MSc. Data Science & Artificial Intelligence

Inverse Problems in Image Processing

Faisal JAYOUSI, Laure BLANC-FERAUD & Luca CALATRONI

Assignment 2

Author: Joris LIMONIER

joris.limonier@gmail.com

Due: March 10, 2023

Contents

1 Exercise 1: Soft-thresholding

The proximal operator of τf is defined as:

$$\operatorname{prox}_{\tau f}(x) = \arg\min_{u \in \mathbb{R}} \frac{1}{2\tau}(u-x)^2 + f(u)$$

which we can apply to the ℓ_1 norm to get:

$$\operatorname{prox}_{\tau|\cdot|}(x) = \arg\min_{u \in \mathbb{R}} \frac{1}{2\tau}(u-x)^2 + |u|$$

Let $h(u)$ be given by:

$$h(u) := \frac{1}{2\tau}(u-x)^2 + |u|$$

The optimality condition states that given that h is is proper, we have:

$$0 \in \partial h(u^*) \iff u^* \in \arg\min_{u \in \mathbb{R}} h(u)$$

then we have:

$$\begin{aligned}\frac{\partial}{\partial u} h(u) &= \frac{\partial}{\partial u}\left[\frac{1}{2\tau}(u-x)^2 + |u|\right] \\ &= \begin{cases} \frac{1}{\tau}(u-x) - 1, & u < 0 \\ \frac{1}{\tau}(u-x) + 1, & u > 0 \end{cases}\end{aligned}$$

We set the derivative to zero to find the critical points of h. We have three cases to consider: $u > 0$, $u < 0$ and $u = 0$.

Case $u > 0$.

$$\begin{aligned}&\frac{\partial}{\partial u} h(u^*) = 0 \\ \implies &\frac{1}{\tau}(u^* - x) + 1 = 0 \\ \implies &u^* = x - \tau\end{aligned}$$

Case $u < 0$.

$$\begin{aligned}&\frac{\partial}{\partial u} h(u^*) = 0 \\ \implies &\frac{1}{\tau}(u^* - x) - 1 = 0 \\ \implies &u^* = x + \tau\end{aligned}$$

Case $u = 0$. In this case, we cannot compute the derivative as the function is non-differentiable in $u = 0$. We have however that the subdifferential of h is given by:

$$\partial h(u) = [-1, 1]$$

In particular, $0 \in \partial h(u)$, so we can apply the optimality condition. Therefore, the proximal operator is given by:

$$\begin{aligned}
\operatorname{prox}_{\tau|\cdot|}(x) &= \begin{cases} x - \tau, & u > 0 \\ x + \tau, & u < 0 \\ 0, & \text{otherwise} \end{cases} \\
&= \begin{cases} x - \tau, & x - \tau > 0 \\ x + \tau, & x + \tau < 0 \\ 0, & \text{otherwise} \end{cases} \\
&= \begin{cases} x - \tau, & x > \tau \\ x + \tau, & x < -\tau \\ 0, & \text{otherwise} \end{cases} \\
&= \begin{cases} x - \tau, & x > \tau \\ x + \tau, & x < -\tau \\ 0, & |x| \leq \tau \end{cases}
\end{aligned}$$

We plot this proximal operator in the companion notebook.

2 Exercise 2: Hard-thresholding

We define f as the ℓ_0 norm:

$$f(x) = |x|_0 = \begin{cases} 0, & x = 0 \\ 1, & x \neq 0 \end{cases}$$

The proximal operator of τf is defined as:

$$\operatorname{prox}_{\tau|\cdot|_0}(x) = \arg\min_{u \in \mathbb{R}} \frac{1}{2\tau} (u - x)^2 + |u|_0$$

Let $h(u)$ be given by:

$$\begin{aligned}
h(u) &:= \frac{1}{2\tau} (u - x)^2 + |u|_0 \\
&= \begin{cases} \frac{1}{2\tau} (u - x)^2 + 1, & u \neq 0 \\ \frac{1}{2\tau} (0 - x)^2 + 0, & u = 0 \end{cases} \\
&= \begin{cases} \frac{1}{2\tau} (u - x)^2 + 1, & u \neq 0 \\ \frac{x^2}{2\tau}, & u = 0 \end{cases}
\end{aligned}$$

We have two cases to consider: $u \neq 0$ and $u = 0$.

Case $u \neq 0$. In this case, h is differentiable. We will compute its derivative and set it to zero to find the critical points of h. The derivative of h is given by:

$$\frac{\partial}{\partial u} h(x) = \frac{1}{\tau}(u - x)$$

We now set this derivative to zero to find the critical points of h:

$$\begin{aligned}
&\frac{\partial}{\partial u} h(u^*) = 0 \\
\implies &\frac{1}{\tau}(u^* - x) = 0 \\
\implies &u^* = x
\end{aligned}$$

Therefore, we have:

$$\begin{aligned}
h(u^*) &= \frac{1}{2\tau}\left(u^* - x\right)^2 + |u^*|_0 \\
&= \frac{1}{2\tau}\left(x - x\right)^2 + 1 \\
&= 1
\end{aligned}$$

Case $u = 0$. In this case, we cannot compute the derivative as the function is non-differentiable in $u = 0$. We have that h is given by:

$$h(u) = \frac{x^2}{2\tau}$$

Now, the question is whether it is better to have $h(u) = \frac{x^2}{2\tau}$ or $h(u) = 1$. We compare the two quantities:

$$\begin{aligned}
\frac{x^2}{2\tau} \leq 1 &\implies x^2 \leq 2\tau \\
&\implies x \in [-\sqrt{2\tau}, \sqrt{2\tau}]
\end{aligned}$$

So in order to minimize h, we prefer having $h(u) = \frac{x^2}{2\tau}$ as long as $x \in [-\sqrt{2\tau}, \sqrt{2\tau}]$ and $h(u) = 1$ otherwise. We need to choose u appropriately, which means that the proximal operator of $\tau|\cdot|_0$ is given by:

$$\operatorname{prox}_{\tau|\cdot|_0}(x) = \begin{cases} 0, & x \in [-\sqrt{2\tau}, \sqrt{2\tau}] \\ x, & \text{otherwise} \end{cases}$$

We plot this proximal operator in the companion notebook.

3 Exercise 3: Non-negativity constraints

3.1 Part 1

Let $\mathbb{R}^n_+$ be the set of vectors with non-negative entries. We define the indicator function of $\mathbb{R}^n_+$ as:

$$\delta_{\mathbb{R}^n_+}(x) = \begin{cases} 0, & x \in \mathbb{R}^n_+ \\ \infty, & \text{otherwise} \end{cases}$$

Therefore the proximal operator of $\delta_{\mathbb{R}^n_+}$ is given by:

$$\text{prox}_{\delta_{\mathbb{R}^n_+}}(x) = \arg\min_{u \in \mathbb{R}^n} \frac{1}{2}\|u - x\|^2 + \delta_{\mathbb{R}^n_+}(u)$$

We define $h(u)$ as:

$$h(u) = \frac{1}{2}\|u - x\|^2 + \delta_{\mathbb{R}^n_+}(u)$$

We understand from the definition of the indicator function that no component of u can be negative, otherwise h would be infinite.
We have two cases to consider: $u \in \mathbb{R}^n_+$ and $u \notin \mathbb{R}^n_+$.

Case $u \in \mathbb{R}^n_+$. In this case, h is differentiable. We will compute its derivative and set it to zero to find the critical points of h. The derivative of h is given by:

$$\begin{aligned} \nabla h(u^*) = 0 &\implies u^* - x = 0 \\ &\implies u^* = x \end{aligned}$$

Case $u \notin \mathbb{R}^n_+$. In this case, as mentioned above, h is infinite. We cannot choose any component u_i of u to be negative, otherwise h would be infinite because of the indicator function. However, we still want to choose u to be as close as possible to x, in order to minimize the $\|u - x\|^2$ component. We proceed in a component-wise fashion, thanks to the separability of h. For all i such that $x_i < 0$, we can therefore choose $u_i = 0$, which is the point in $\mathbb{R}_+$ which minimizes the distance to x_i.
We can summarize the proximal operator of $\delta_{\mathbb{R}^n_+}$ as:

$$\forall i = 1, \dots, n, \ \left\{\text{prox}_{\delta_{\mathbb{R}^n_+}}(x)\right\}_i = \max(0, x_i)$$

3.2 Part 2

We now compute the proximal operator of $\tau|\cdot|_1 + \delta_{\mathbb{R}^n_+}(\cdot)$, which is given by:

$$\text{prox}_{\tau|\cdot|_1 + \delta_{\mathbb{R}^n_+}(\cdot)}(x) = \arg\min_{u \in \mathbb{R}^n_+} \frac{1}{2\tau}\|u - x\|^2 + |u|_1 + \delta_{\mathbb{R}^n_+}(u)$$

We define $h(u)$ as:

$$h(u) = \frac{1}{2\tau}\|u - x\|^2 + |u|_1 + \delta_{\mathbb{R}^n_+}(u)$$

We note that h is separable. We can therefore apply the proximal operator of $\tau|\cdot|_1$ to each component of u. Moreover, we can use the same argument as in section 3.1 regarding the non-negativity constraint. Indeed once again, as long as one of the component of u is negative, h is infinite, so minimizing h means that all components of u must be non-negative. For the rest, the minimizer of $\frac{1}{2\tau}\|u-x\|^2 + |u|_1$ is by definition equal to $\text{prox}_{\tau|\cdot|_1}(x)$, that is:

$$\arg\min_{u\in\mathbb{R}^n} \frac{1}{2\tau}\|u-x\|^2 + |u|_1 =: \text{prox}_{\tau|\cdot|_1}(x)$$

and we computed this proximal operator in section 1.
Now, the indicator forces all components of u to be non-negative (this is the same argument as in section 3.1 really). As a result, we can summarize the proximal operator of $\tau|\cdot|_1 + \delta_{\mathbb{R}^n_+}(\cdot)$ as:

$$\forall i = 1,\dots,n, \ \left(\text{prox}_{\tau|\cdot|_1+\delta_{\mathbb{R}^n_+}(\cdot)}(x)\right)_i = \max\left(\left(\text{prox}_{\tau|\cdot|_1}(x)\right)_i, 0\right)$$

where the maximum is taken component-wise.

4 Exercise 4

Let us define f as the elastic net functional, that is:

$$f(x) = \|x\|_1 + \frac{\lambda}{2}\|x\|^2$$

We want ot compute $\text{prox}_f(x)$.
We define g as:

$$g(x) = \|x\|_1$$

therefore, we have that:

$$f(x) = g(x) + \frac{\lambda}{2}\|x\|^2$$

By proposition 3 with $c = \lambda$, we have that:

$$\text{prox}_f(x) = \text{prox}_{\frac{g}{\lambda+1}}\left(\frac{x}{\lambda+1}\right)$$

Now, using proposition 2 with the variable from the proposition λ_{prop} being such that $\lambda_{\text{prop}} = \lambda + 1$, we have that:

$$\begin{aligned}
&(\lambda+1)\,\text{prox}_{\frac{g}{\lambda+1}}\left(\frac{x}{\lambda+1}\right) = \text{prox}_g(x)\\
\implies &\text{prox}_{\frac{g}{\lambda+1}}\left(\frac{x}{\lambda+1}\right) = \frac{1}{\lambda+1}\text{prox}_g(x)\\
\implies &\text{prox}_f(x) = \frac{1}{\lambda+1}\text{prox}_g(x)\\
\implies &\text{prox}_f(x) = \frac{1}{\lambda+1}\text{prox}_{|\cdot|_1}(x)
\end{aligned}$$

where we know section 1 the value of the proximal operator of $|\cdot|_1$.