Example Solutions for Assignment 7

Problem 1 (The Splitting Validation)
In order to show that the Douglas-Rachford Splitting $J^\lambda_{DR}(u) := [J^\lambda_{\partial R}(2J^\lambda_{\partial S} - I) + (I - J^\lambda_{\partial S})]u$ is a valid splitting scheme, we have to show $0 \in J_F(u) \Leftrightarrow u = J^\lambda_{DR}(u)$

$$\begin{aligned}
& u && = J^\lambda_{DR}(u) \\
\Leftrightarrow\quad & u && = [J^\lambda_{\partial R}(2J^\lambda_{\partial S} - I) + (I - J^\lambda_{\partial S})]u \\
\Leftrightarrow\quad & u && = 2J^\lambda_{\partial R}J^\lambda_{\partial S}u - J^\lambda_{\partial R}u + u - J^\lambda_{\partial S}u \\
\Leftrightarrow\quad & u && = 2(I + \lambda\partial R)^{-1}(I + \lambda\partial S)^{-1}u + (I + \lambda\partial R)^{-1}u \\
& && \quad + u - (I + \lambda\partial S)^{-1}u \\
\Leftrightarrow\quad & (I + \lambda\partial R)(I + \lambda\partial S)u && = 2u - (I + \lambda\partial S)u \\
& && \quad + (I + \lambda\partial R)(I + \lambda\partial S)u - (I + \lambda\partial R)u \\
\Leftrightarrow\quad & 0 && = 2u - u - \lambda\partial S - u - \lambda\partial Ru \\
\Leftrightarrow\quad & -u && = -u - \lambda Su - \lambda Ru \\
\Leftrightarrow\quad & u && = (I + \lambda\underbrace{(\partial S + \partial R)}_{\partial F})u \\
\Leftrightarrow\quad & u && = (I + \lambda\partial F)u \\
\Leftrightarrow\quad & (I + \lambda\partial F)^{-1}u && = u \\
\Leftrightarrow\quad & u && = J^\lambda_{\partial F}(u) \\
\Leftrightarrow\quad & 0 && \in J^\lambda_F(u)
\end{aligned}$$

Problem 2 (The Conjugate Convexification)
Our task is to find the convex conjugate function

$$f^*(y) \quad = \quad \sup_{x \in D} y^\top x - f(x).$$

1. $f(x) = \exp(x)$:

$$f^*(y) \quad = \quad \sup_{x \in D} xy - \exp(x).$$

   For a supremum, the derivative of f^* w.r.t. x should be zero, i.e. $(f^*)'(x) = y - \exp(x) \overset{!}{=} 0$. From this it follows $x = \log y$. Furthermore, $(f^*)''(x) = -\exp(x) < 0$, so this is indeed a supremum. By plugging this result back into the definition of $f^*(y)$ and checking some cases, we arrive at

$$f^*(y) = \begin{cases} \log(y)y - y & (y > 0) \\ 0 & (y = 0) \\ \infty & (y < 0) \end{cases}$$

2. $f(x) = |x|$:

$$f^*(y) = \sup_{x \in D} xy - |x|.$$

Again, by deriving f^* w.r.t. x, we arrive at the necessary condition for a supremum: $y - \mathrm{sgn}(x) \overset{!}{=} 0$. Plugging this back into f^*, we arrive at the following result:

$$f^*(y) = \begin{cases} 0 & (|y| \leq 1) \\ \infty & (|y| > 1) \end{cases}$$

3. $f(x) = \frac{1}{2}x^2$:

$$f^*(y) = \sup_{x \in D} xy - \frac{1}{2}x^2.$$

After deriving f^* we arrive at $(f^*)'(x) = y - x \overset{!}{=} 0$. Plugging this back into f^*, we get

$$f^*(y) = \frac{1}{2}y^2.$$

4. $f(x) = a^\top x - b = x^\top a - b$:

$$f^*(y) = \sup_{x \in D} x^\top y - x^\top a + b.$$

Deriving w.r.t. to x we arrive at $(f^*)'(x) = y - a \overset{!}{=} 0$. Then the resulting convex conjugate function is

$$f^*(y) = \begin{cases} b & (y = a) \\ \infty & (y \neq a) \end{cases}$$

Problem 3 (Musings on Bregman Distance)

1. For the non-negativity we make use of the proper convex function property, i.e. $f(x) = x \cdot b - \beta$ as in (13.9) . We consider

$$\begin{aligned} B_F(p,q) &= F(p) - F(q) - (p-q) \cdot \nabla F(q) \geq 0 \\ &\Leftrightarrow p \cdot b - \beta - q \cdot b + \beta - (p-q) \cdot b \geq 0 \\ &\Leftrightarrow (p-q) \cdot b \geq (p-q) \cdot b, \end{aligned}$$

which holds for any arbitrary b.

Another, more general method is to use a Taylor expansion for $F(p)$ around position q:

$$\begin{aligned} F(p) = F(q) + (p-q)^\top \nabla F(q) + \frac{1}{2}(p-q)^\top HF(q)(p-q) + \mathcal{O}(q^3) \\ \Leftrightarrow \underbrace{F(p) - F(q) - (p-q)^\top \nabla F(q)}_{=B_F(p,q)} = \frac{1}{2}(p-q)^\top \underbrace{HF(q)}_{\geq 0}(p-q) + \mathcal{O}(q^3) \end{aligned}$$

As we can see the left hand side of the equation is the Bregman distance and on the right hand side, we have a second order term. We have chosen F to be convex for all p, q, therefore, the Hessian on the right hand side is positive definite, i.e. the right hand side will become bigger than zero for $q \to 0$.

2. We only show this for the 1-D case. In order to assure convexity of the Bregman distance, we have to assure that the Hessian of the Bregman distance function is positive semi-definite. One way to check this is to consider the main minors of the Hessian. Let us now compute the derivatives.

$$\begin{aligned} \frac{\partial}{\partial p} B_F(p,q) &= F'(p) - F'(q) \\ \frac{\partial}{\partial q} B_F(p,q) &= -(p-q)F''(q) \\ \frac{\partial^2}{\partial p^2} B_F(p,q) &= F''(p) \\ \frac{\partial^2}{\partial p \partial q} B_F(p,q) &= -F''(q) \\ \frac{\partial^2}{\partial q^2} B_F(p,q) &= -(p-q)F'''(q) + F''(q) \end{aligned}$$

A good way to check positive semi-definiteness with the main minors. All determinants of the main minors have to be positive in order to assure positive semi-definiteness. The first minor $F''(p)$ should be positive. This means that the distance function is at least convex in the first argument. However, the determinant for the second minor is given by $F''(p)(F''(q) - (p-q)F'''(q)) - (F''(q))^2$. Unfortunately, we cannot state anything here now for q, so it may or may not be F to be convex in q. Hence, resulting in the statement described. $B_F(p,q)$ is convex in its first argument, but not necessarily in its second.

3.

$$\begin{aligned}
B_{F_1+\lambda F_2}(p,q) &= (F_1+\lambda F_2)(p) - (F_1+\lambda F_2)(q) - (p-q)\nabla(F_1+\lambda F_2)(q) \\
&= F_1(p) + \lambda F_2(p) - F_1(q) - \lambda F_2(q) - (p-q)\nabla F_1(q) - (p-q)\nabla\lambda F_2(q) \\
&= B_{F_1}(p,q) + \lambda B_{F_2}(p,q)
\end{aligned}$$

Problem 4 (The Diverging Bregman)
At first, we consider the derivative of a function $f(x) = x\log x - x$, i.e. $f'(x) = \log x$. Applied on our given function, this gives $\nabla F(q) = (\log q_1, \log q_2, \ldots, \log q_n)^\top$. This gives us

$$\begin{aligned}
B_F(p,q) &= F(p) - F(q) - (p-q)\cdot\nabla F(q) \\
&= \sum_i p_i \log p_i - \sum_i p_i - \sum_i q_i \log q_i + \sum_i q_i - \sum_i (p_i - q_i)\log q_i \\
&= \sum p_i \log\frac{p_i}{q_i} - \sum_i p_i + \sum_i q_i.
\end{aligned}$$

If we suppose now, that $\sum_i p_i = \sum_i q_i = 1$, this results in the so-called Kullback-Leibler divergence.

$$B_{KL}(p,q) = \sum p_i \log\frac{p_i}{q_i}.$$

Problem 5 (The ROF Lagrangian)
In this exercise we want to compute a PDE for the Ruder-Osher-Fatemi model. From the variational model

$$\int_\Omega \|\nabla u\| + \frac{\lambda}{2}\|u-f\|^2 \, \mathrm{d}x \, \mathrm{d}y$$

From this, we get the Lagrangian

$$F(x,y,u,u_x,u_y) = \sqrt{u_x^2 + u_y^2} + \frac{\lambda}{2}(u-f)^2$$

For the Euler-Lagrange equation

$$F_u - \frac{\mathrm{d}}{\mathrm{d}x} F_{u_x} - \frac{\mathrm{d}}{\mathrm{d}y} F_{u_y}$$

with the ingredients

$$\begin{aligned} F_u &= \lambda(u - f) \\ F_{u_x} &= \frac{u_x}{\sqrt{u_x^2 + u_y^2}} \\ F_{u_y} &= \frac{u_y}{\sqrt{u_x^2 + u_y^2}} \end{aligned}$$

we arrive at the PDE

$$\lambda(u - f) - \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|}\right)$$

Obviously this PDE is not differentiable at positions where $u_x = u_y = 0$. This problem is mostly being solved by artificially adding a small number ε in the norm, i.e. $\|\nabla u\|_\varepsilon := \sqrt{u_x^2 + u_y^2 + \varepsilon^2}$, resulting in

$$\lambda(u - f) - \operatorname{div}\left(\frac{\nabla u}{\|\nabla u\|_\varepsilon}\right)$$
