Overlapping Group Sparsity (OGS) Equations

Ivan Selesnick

Last edit: May 25, 2016

May 13, 2016

1 Introduction

This note explains equations (16) and (17) in the paper [1] on denoising of group-sparse signals.

For simplicity, in these notes we set the group size to 2, i.e., $K = 2$. Let the signal x be a sequence,

$$x = \ \dots,\ x_{-1},\ x_0,\ x_1,\ x_2,\ x_3,\ \dots \tag{1}$$

with $x_i \in \mathbb{R}$. Then the regularizer (penalty function) in the paper is given by

$$R(x) = \cdots + \sqrt{x_1^2 + x_2^2} + \sqrt{x_2^2 + x_3^2} + \sqrt{x_3^2 + x_4^2} + \cdots \tag{2}$$

Our goal is to find a quadratic separable majorizer which we denote R^{M}. This function should satisfy

$$R^{\mathsf{M}}(x; u) \geqslant R(x), \quad R^{\mathsf{M}}(u; u) = R(u). \tag{3}$$

Note that for $t \in \mathbb{R}$ and $v \in \mathbb{R}$,

$$\frac{1}{2}\frac{t^2}{|v|} + \frac{1}{2}|v| \geqslant |t| \tag{4}$$

for all $v \neq 0$, as illustrated in Fig. 1. The figure shows that the left-hand side of (4) is an upper bound of the absolute value function, and the two functions touch at $t = v$.

Since (4) is satisfied for any $v \neq 0$, we can write

$$\frac{1}{2}\frac{x_1^2 + x_2^2}{\sqrt{u_1^2 + u_2^2}} + \frac{1}{2}\sqrt{u_1^2 + u_2^2} \geqslant \sqrt{x_1^2 + x_2^2} \tag{5}$$

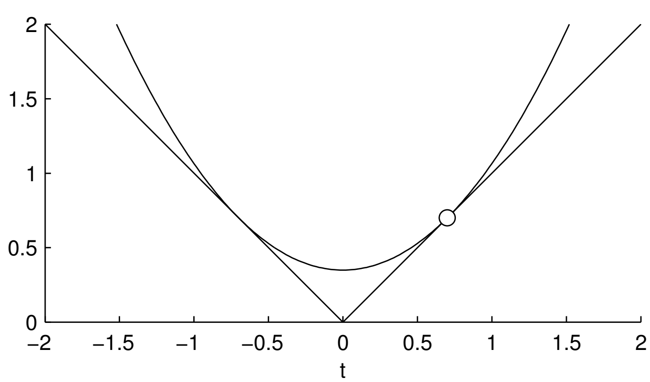

Figure 1: Illustration of (4) for $u = 0.7$ where the two functions touch.

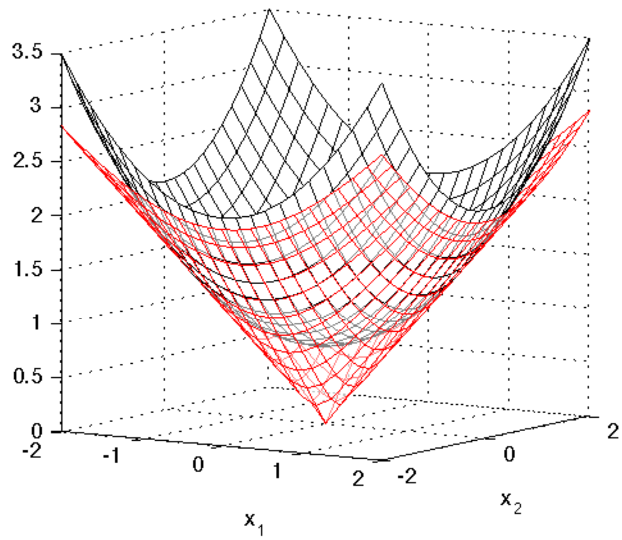

Figure 2: Illustration of (5) for $(u_1, u_2) = (0.7, 0.3)$.

for all (u_1, u_2) such that $u_1^2 + u_2^2 > 0$. Equation (5) is illustrated in Fig. 2. The function $\sqrt{x_1^2 + x_2^2}$ shown in red is the lower bound. The function shown in black is the the quadratic majorizer.

A majorizer of R in (2) is obtained by using (5) for each term, leading to

$$R^{\mathsf{M}}(x; u) = \cdots + \frac{1}{2}\frac{x_1^2 + x_2^2}{\sqrt{u_1^2 + u_2^2}} + \frac{1}{2}\sqrt{u_1^2 + u_2^2} + \frac{1}{2}\frac{x_2^2 + x_3^2}{\sqrt{u_2^2 + u_3^2}} + \frac{1}{2}\sqrt{u_2^2 + u_3^2} + \frac{1}{2}\frac{x_3^2 + x_4^2}{\sqrt{u_3^2 + u_4^2}} + \frac{1}{2}\sqrt{u_3^2 + u_4^2} + \cdots \tag{6}$$

Collecting like terms x_i^2, we write

$$R^{\mathsf{M}}(x; u) = \cdots + \frac{1}{2}x_1^2\left(\frac{1}{\sqrt{u_0^2 + u_1^2}} + \frac{1}{\sqrt{u_1^2 + u_2^2}}\right) + \frac{1}{2}x_2^2\left(\frac{1}{\sqrt{u_1^2 + u_2^2}} + \frac{1}{\sqrt{u_2^2 + u_3^2}}\right) \frac{1}{2}x_3^2\left(\frac{1}{\sqrt{u_2^2 + u_3^2}} + \frac{1}{\sqrt{u_3^2 + u_4^2}}\right) + \cdots \cdots + \frac{1}{2}\sqrt{u_1^2 + u_2^2} + \frac{1}{2}\sqrt{u_2^2 + u_3^2} + \frac{1}{2}\sqrt{u_3^2 + u_4^2} \cdots \tag{7}$$

That is,

$$R^{\mathsf{M}}(x; u) = \frac{1}{2}\left(\cdots + r(1, u)\, x_1^2 + r(2, u)\, x_2^2 + r(3, u)\, x_3^2 + \cdots\right) + C(u) \tag{8}$$

where $C(u)$ does not depend on any x_i and

$$r(1, u) = \frac{1}{\sqrt{u_0^2 + u_1^2}} + \frac{1}{\sqrt{u_1^2 + u_2^2}} \tag{9}$$

$$r(2, u) = \frac{1}{\sqrt{u_1^2 + u_2^2}} + \frac{1}{\sqrt{u_2^2 + u_3^2}} \tag{10}$$

$$r(3, u) = \frac{1}{\sqrt{u_2^2 + u_3^2}} + \frac{1}{\sqrt{u_3^2 + u_4^2}}, \tag{11}$$

etc. These are equations (16) and (17) in the paper [1] for the special case of $K = 2$. For $K > 2$, the situation is similar.

The objective function F considered in the paper [1] is defined by

$$F(x) = \frac{1}{2}\|y - x\|_2^2 + \lambda R(x) \tag{12}$$

which, for $K = 2$, is expanded as

$$F(x) = \frac{1}{2}\|y - x\|_2^2 + \lambda \left(\cdots + \sqrt{x_1^2 + x_2^2} + \sqrt{x_2^2 + x_3^2} + \sqrt{x_3^2 + x_4^2} + \cdots \right). \tag{13}$$

A majorizer of F is therefore given by

$$F(x) = \frac{1}{2}\|y - x\|_2^2 + \lambda R^{\mathsf{M}}(x; u) \tag{14}$$

which is the same as

$$F^{\mathsf{M}}(x; u) = \frac{1}{2}\|y - x\|_2^2 + \frac{\lambda}{2}\Big(\cdots + r(1, u)\, x_1^2 + r(2, u)\, x_2^2 + r(3, u)\, x_3^2 + \cdots \Big) + \lambda\, C(u) \tag{15}$$

or

$$F^{\mathsf{M}}(x; u) = \frac{1}{2} \sum_i \left[(y_i - x_i)^2 + \lambda\, r(i, u)\, x_i^2 \right] + \lambda\, C(u) \tag{16}$$

The function F^{M} is separable (additive) in x, so the optimal x_i can be found by minimizing with respect to each x_i independently. Minimizing $F^{\mathsf{M}}(x; u)$ with respect to x_i is equivalent to minimizing

$$f(x_i) = \frac{1}{2} \left[(y_i - x_i)^2 + \lambda\, r(i, u)\, x_i^2 \right]. \tag{17}$$

The derivative is

$$f'(x_i) = x_i - y_i + \lambda\, r(i, u)\, x_i. \tag{18}$$

Setting the derivative to zero and solving for x_i yields the minimizer of $F^{\mathsf{M}}(x; u)$ to be

$$x_i = \frac{y_i}{1 + \lambda\, r(i, u)}. \tag{19}$$

The Majorization-Minimization (MM) optimization method comprises the iteration

$$x^{(k+1)} = \arg\min_x F^{\mathsf{M}}(x, x^{(k)}). \tag{20}$$

Hence, we obtain the update equation

$$x_i^{(k+1)} = \frac{y_i}{1 + \lambda\, r(i, x^{(k)})} \tag{21}$$

which is equation (23) in the paper [1].

References

[1] P.-Y. Chen and I. W. Selesnick. Translation-invariant shrinkage/thresholding of group sparse signals. *Signal Processing*, 94:476–489, January 2014.