MA3111: Mathematical Image Processing
Alternating Direction Method of Multipliers

Suh-Yuh Yang (楊肅煜)

Department of Mathematics, National Central University
Jhongli District, Taoyuan City 320317, Taiwan

First version: October 24, 2021/Last updated: June 24, 2025

Outline of "alternating direction method of multipliers"

- This lecture will briefly introduce the alternating direction method of multipliers (ADMM) for solving linearly equality-constrained minimization problems.
- The material of this lecture is based on the long review paper:

  S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Distributed optimization and statistical learning via the ADMM, *Foundations and Trends in Machine Learning, 3 (2010), pp. 1-122.*

The sparse representation problem

We consider the following minimization problem:

Sparse representation problem: *Given a signal vector $\boldsymbol{x} \in \mathbb{R}^m$ and a dictionary matrix $\boldsymbol{D} \in \mathbb{R}^{m \times n}$ with $m < n$, we seek a sparse coefficient vector $\boldsymbol{z}^* \in \mathbb{R}^n$ for a fixed parameter $\lambda > 0$ such that*

$$\boldsymbol{z}^* = (\in) \arg\min_{\boldsymbol{z} \in \mathbb{R}^n} F(\boldsymbol{z}) := \frac{1}{2} \|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2 + \lambda \|\boldsymbol{z}\|_1 . \qquad (\star)$$

Properties of function F:

- F is a continuous function, $F(\boldsymbol{z}) \geq 0 \ \forall \ \boldsymbol{z} \in \mathbb{R}^n$, and F is coercive since $\lim_{\|\boldsymbol{z}\| \to \infty} F(\boldsymbol{z}) = \infty$. Here, $\| \cdot \|$ can be arbitrary vector norm on $\mathbb{R}^n$ due to the norm-equivalence for the finite dimensional vector space $\mathbb{R}^n$.
- Let $f(\boldsymbol{z}) := \|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2$. Then we have a quadratic function

  $$f(\boldsymbol{z}) = \langle \boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}, \boldsymbol{x} - \boldsymbol{D}\boldsymbol{z} \rangle = \boldsymbol{z}^\top \boldsymbol{D}^\top \boldsymbol{D}\boldsymbol{z} + 2(-\boldsymbol{x}^\top \boldsymbol{D})\boldsymbol{z} + \|\boldsymbol{x}\|_2^2,$$

  where $\boldsymbol{A} := \boldsymbol{D}^\top \boldsymbol{D}$ is symmetric and positive semidefinite ($\boldsymbol{A} \succeq \boldsymbol{0}$). Thus, f is a convex function. Since $g(\boldsymbol{z}) := \|\boldsymbol{z}\|_1$ is also a convex function, $F(\boldsymbol{z}) = \frac{1}{2} f(\boldsymbol{z}) + \lambda g(\boldsymbol{z})$ is therefore convex over $\mathbb{R}^n$.

The existence and uniqueness of solution of problem $(\star)$

- **Existence:** *Since F is continuous and coercive on $\mathbb{R}^n$, the sparse representation problem $(\star)$ has a global minimum point in $\mathbb{R}^n$, and the set of all global minimizers is convex.*
- **No uniqueness:** *In general, the solution of problem $(\star)$ may not be unique when $m < n$, even if the matrix $\boldsymbol{D}$ is of full rank. If $m < n$ and $\operatorname{rank}(\boldsymbol{D}) = m$, then the $n \times n$ matrix $\boldsymbol{A} := \boldsymbol{D}^\top \boldsymbol{D}$ is symmetric and $\operatorname{rank}(\boldsymbol{A}) = m$. Hence, $\boldsymbol{A}$ is not invertible and $\boldsymbol{A} \not\succ \boldsymbol{0}$. As a result, F is not strictly convex.*

  Certain additional conditions can guarantee the uniqueness of the solution to problem $(\star)$.
- Problem $(\star)$ is also a regression analysis method in statistics and machine learning. It is the so-called *least absolute shrinkage and selection operator (LASSO).*

  R. J. Tibshirani, The lasso problem and uniqueness, *Electronic Journal of Statistics*, 7 (2013), pp. 1456-1490 $\oplus$ A. Ali, 13 (2019), pp. 2307-2347.

Alternating direction method of multipliers

We will use the *"alternating direction method of multipliers (ADMM)"* to solve the above ℓ_1-norm sparse representation problem.

- ADMM is an iterative scheme for solving the following equality constrained convex/nonconvex optimization problems:

$$\min_{\boldsymbol{z}} f(\boldsymbol{z}) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{z} = \boldsymbol{b}.$$

- ADMM consists of three steps:
  (1) *adding an auxiliary variable* $\boldsymbol{y}$ *and a dual variable (multipliers)* $\boldsymbol{v}$ *and then scaled as* $\boldsymbol{u}$
  (2) *separating the new cost function into a sum of* $f(\boldsymbol{z})$ *and* $g(\boldsymbol{y})$
  (3) *using an iterative method to solve the problem*

- Then the optimization problem can be re-posed as

$$\min_{\boldsymbol{z},\,\boldsymbol{y}} \big(f(\boldsymbol{z}) + g(\boldsymbol{y})\big) \quad \text{subject to} \quad \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} = \boldsymbol{c}.$$

Derivation of the ADMM: augmented Lagrangian function

First, we formulate the *augmented Lagrangian function*

$$L_\rho(\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}) := f(\boldsymbol{z}) + g(\boldsymbol{y}) + \underbrace{\boldsymbol{v}^\top}_{\text{multipliers}} (\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c}) + \underbrace{\frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c}\|_2^2}_{\text{penalty term}},$$

where $\rho > 0$ is the penalty parameter. Then the iterative scheme of the *augmented Lagrangian method (ALM)* is given by

$$\begin{aligned}
(\boldsymbol{z}^{(i+1)}, \boldsymbol{y}^{(i+1)}) &= \arg\min_{\boldsymbol{z},\boldsymbol{y}} L_\rho(\boldsymbol{z},\boldsymbol{y},\boldsymbol{v}^{(i)}), \\
\boldsymbol{v}^{(i+1)} &= \boldsymbol{v}^{(i)} + \rho\big(\boldsymbol{A}\boldsymbol{z}^{(i+1)} + \boldsymbol{B}\boldsymbol{y}^{(i+1)} - \boldsymbol{c}\big),
\end{aligned}$$

where the second equation is obtained by the *dual ascent method.*

In ADMM, $\boldsymbol{z}$ and $\boldsymbol{y}$ are updated in an alternating or sequential fashion, which accounts for the term *alternating direction*.

$$\begin{aligned}
\boldsymbol{z}^{(i+1)} &= \arg\min L_\rho(\boldsymbol{z},\boldsymbol{y}^{(i)},\boldsymbol{v}^{(i)}), \\
\boldsymbol{y}^{(i+1)} &= \arg\min L_\rho(\boldsymbol{z}^{(i+1)},\boldsymbol{y},\boldsymbol{v}^{(i)}), \\
\boldsymbol{v}^{(i+1)} &= \boldsymbol{v}^{(i)} + \rho(\boldsymbol{A}\boldsymbol{z}^{(i+1)} + \boldsymbol{B}\boldsymbol{y}^{(i+1)} - \boldsymbol{c}).
\end{aligned}$$

Scaled form of the augmented Lagrangian

The ADMM can be written in a slightly different form, which is often more convenient, by combining the linear and quadratic terms in the augmented Lagrangian and scaling the dual variable (multipliers) $\boldsymbol{v}$.

Define the residual $\boldsymbol{r} := \boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c}$. Then

$$\begin{aligned}
&\boldsymbol{v}^\top(\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c}\|_2^2 \\
&\qquad = \boldsymbol{v}^\top\boldsymbol{r} + \frac{\rho}{2}\|\boldsymbol{r}\|_2^2 = \frac{\rho}{2}\|\boldsymbol{r} + \frac{1}{\rho}\boldsymbol{v}\|_2^2 - \frac{1}{2\rho}\|\boldsymbol{v}\|_2^2.
\end{aligned}$$

Set $\boldsymbol{u} = \frac{1}{\rho}\boldsymbol{v}$. Then $L_\rho(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{v}) = L_\rho(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{u})$, and

$$L_\rho(\boldsymbol{z}, \boldsymbol{y}, \boldsymbol{u}) = f(\boldsymbol{z}) + g(\boldsymbol{y}) + \frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2.$$

ADMM: scaled form

The ADMM in the scaled form is given by

$$\begin{aligned}
\boldsymbol{z}^{(i+1)} &= \arg\min_{\boldsymbol{z}} \Big(f(\boldsymbol{z}) + g(\boldsymbol{y}^{(i)}) + \frac{\rho}{2} \|\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y}^{(i)} - \boldsymbol{c} + \boldsymbol{u}^{(i)}\|_2^2 - \frac{\rho}{2} \|\boldsymbol{u}^{(i)}\|_2^2 \Big), \\
\boldsymbol{y}^{(i+1)} &= \arg\min_{\boldsymbol{y}} \Big(f(\boldsymbol{z}^{(i+1)}) + g(\boldsymbol{y}) + \frac{\rho}{2} \|\boldsymbol{A}\boldsymbol{z}^{(i+1)} + \boldsymbol{B}\boldsymbol{y} - \boldsymbol{c} + \boldsymbol{u}^{(i)}\|_2^2 - \frac{\rho}{2} \|\boldsymbol{u}^{(i)}\|_2^2 \Big), \\
\boldsymbol{u}^{(i+1)} &= \boldsymbol{u}^{(i)} + \boldsymbol{A}\boldsymbol{z}^{(i+1)} + \boldsymbol{B}\boldsymbol{y}^{(i+1)} - \boldsymbol{c},
\end{aligned}$$

where $\rho > 0$ is the *penalty parameter* which is related to the convergent rate of the iterations.

Note that the terms in blue can be omitted in practical computations!

Some convergence analysis can be found in the following paper:

S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, Distributed optimization and statistical learning via the ADMM, *Foundations and Trends in Machine Learning*, 3 (2010), pp. 1-122.

ADMM for the ℓ_1-norm sparse representation problem

- For the ℓ_1-norm sparse representation problem,

$$\boldsymbol{z}^* = \arg\min_{\boldsymbol{z}\in\mathbb{R}^n} \left(\frac{1}{2} \|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2 + \lambda \|\boldsymbol{z}\|_1 \right), \qquad \lambda > 0, \qquad (\star)$$

we set

$$f(\boldsymbol{z}) := \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2,$$
$$g(\boldsymbol{y}) := \lambda\|\boldsymbol{y}\|_1,$$
$$\boldsymbol{z} - \boldsymbol{y} = \boldsymbol{0}. \qquad (\boldsymbol{A}\boldsymbol{z} + \boldsymbol{B}\boldsymbol{y} = \boldsymbol{c})$$

- The ADMM for the problem is given by

$$\boldsymbol{z}^{(i+1)} = \arg\min_{\boldsymbol{z}} \left(\frac{1}{2}\|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2 + \frac{\rho}{2}\|\boldsymbol{z} - \boldsymbol{y}^{(i)} + \boldsymbol{u}^{(i)}\|_2^2 \right), \qquad (P_1)$$

$$\boldsymbol{y}^{(i+1)} = \arg\min_{\boldsymbol{y}} \left(\lambda\|\boldsymbol{y}\|_1 + \frac{\rho}{2}\|\boldsymbol{z}^{(i+1)} - \boldsymbol{y} + \boldsymbol{u}^{(i)}\|_2^2 \right), \qquad (P_2)$$

$$\boldsymbol{u}^{(i+1)} = \boldsymbol{u}^{(i)} + \boldsymbol{z}^{(i+1)} - \boldsymbol{y}^{(i+1)}, \qquad (P_3)$$

where $\rho > 0$ is *penalty parameter* related to the convergent rate.

Solving minimization problem (P_1)

Define the function

$$F_1(\boldsymbol{z}) := \frac{1}{2}\|\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}\|_2^2 + \frac{\rho}{2}\|\boldsymbol{z} - \boldsymbol{y}^{(i)} + \boldsymbol{u}^{(i)}\|_2^2.$$

Then F_1 is a quadratic function in variables $z_1, z_2, \cdots, z_n$ and $F_1(\boldsymbol{z}) \geq 0 \; \forall \; \boldsymbol{z} \in \mathbb{R}^n$. To solve "$\min_{\boldsymbol{z}} F_1(\boldsymbol{z})$", first we compute

$$\begin{aligned} \nabla F_1(\boldsymbol{z}) &= -\boldsymbol{D}^\top(\boldsymbol{x} - \boldsymbol{D}\boldsymbol{z}) + \rho\boldsymbol{I}(\boldsymbol{z} - \boldsymbol{y}^{(i)} + \boldsymbol{u}^{(i)}) \\ &= (\boldsymbol{D}^\top\boldsymbol{D} + \rho\boldsymbol{I})\boldsymbol{z} - \left(\boldsymbol{D}^\top\boldsymbol{x} + \rho(\boldsymbol{y}^{(i)} - \boldsymbol{u}^{(i)})\right). \end{aligned}$$

Letting $\nabla F_1(\boldsymbol{z}) = \mathbf{0}$, we have the linear system

$$(\boldsymbol{D}^\top\boldsymbol{D} + \rho\boldsymbol{I})\boldsymbol{z} = \left(\boldsymbol{D}^\top\boldsymbol{x} + \rho(\boldsymbol{y}^{(i)} - \boldsymbol{u}^{(i)})\right), \qquad (\star\star)$$

and mathematically,

$$\boldsymbol{z}^{(i+1)} = (\boldsymbol{D}^\top\boldsymbol{D} + \rho\boldsymbol{I})^{-1}\left(\boldsymbol{D}^\top\boldsymbol{x} + \rho(\boldsymbol{y}^{(i)} - \boldsymbol{u}^{(i)})\right).$$

Notice that $\boldsymbol{D}^\top\boldsymbol{D} + \rho\boldsymbol{I}$ is SPD and then $(\star\star)$ can be numerically solved efficiently.

Solving minimization problem (P_2)

The solution of problem (P_2),

$$\boldsymbol{y}^{(i+1)} = \arg\min_{\boldsymbol{y}} \Big(\lambda \|\boldsymbol{y}\|_1 + \frac{\rho}{2} \|\boldsymbol{z}^{(i+1)} - \boldsymbol{y} + \boldsymbol{u}^{(i)}\|_2^2 \Big),$$

has the closed form (see next few pages):

$$\boldsymbol{y}^{(i+1)} = \mathcal{S}_{\lambda/\rho}(\boldsymbol{z}^{(i+1)} + \boldsymbol{u}^{(i)}),$$

where *the soft-thresholding (軟閾值) function* $\mathcal{S}_{\lambda/\rho}$, is defined by

$$\mathcal{S}_{\lambda/\rho}(\boldsymbol{v}) := \operatorname{sign}(\boldsymbol{v}) \odot \max(\boldsymbol{0},\ |\boldsymbol{v}| - \lambda/\rho),$$

and $\operatorname{sign}(\cdot)$, $\max(\cdot,\cdot)$, and $|\cdot|$ are all applied to the input vector $\boldsymbol{v}$ component-wisely, and $\odot$ is the Hadamard product.

Finally, the iterative scheme can be posed as follows:

$$\begin{aligned}
\boldsymbol{z}^{(i+1)} &= (\boldsymbol{D}^\top \boldsymbol{D} + \rho \boldsymbol{I})^{-1} (\boldsymbol{D}^\top \boldsymbol{x} + \rho(\boldsymbol{y}^{(i)} - \boldsymbol{u}^{(i)})),\\
\boldsymbol{y}^{(i+1)} &= \mathcal{S}_{\lambda/\rho}(\boldsymbol{z}^{(i+1)} + \boldsymbol{u}^{(i)}),\\
\boldsymbol{u}^{(i+1)} &= \boldsymbol{u}^{(i)} + \boldsymbol{z}^{(i+1)} - \boldsymbol{y}^{(i+1)}.
\end{aligned}$$

Details of the solution of problem (P_2)

Recall the problem (P_2),

$$\boldsymbol{y}^{(i+1)} = \arg\min_{\boldsymbol{y}} \left(\lambda \|\boldsymbol{y}\|_1 + \frac{\rho}{2} \|\boldsymbol{z}^{(i+1)} - \boldsymbol{y} + \boldsymbol{u}^{(i)}\|_2^2 \right). \qquad (P_2)$$

Let $\boldsymbol{v} := \boldsymbol{z}^{(i+1)} + \boldsymbol{u}^{(i)} \in \mathbb{R}^n$. Then we have

$$\boldsymbol{y}^{(i+1)} = \arg\min_{\boldsymbol{y}} \left(\lambda \|\boldsymbol{y}\|_1 + \frac{\rho}{2} \|\boldsymbol{v} - \boldsymbol{y}\|_2^2 \right).$$

Define a real-valued function $F_2(\boldsymbol{y})$ as follows:

$$\begin{aligned} F_2(\boldsymbol{y}) &= \lambda \|\boldsymbol{y}\|_1 + \frac{\rho}{2} \|\boldsymbol{v} - \boldsymbol{y}\|_2^2 \\ &= \left(\lambda |y_1| + \frac{\rho}{2}(v_1 - y_1)^2 \right) + \cdots + \left(\lambda |y_n| + \frac{\rho}{2}(v_n - y_n)^2 \right) \\ &:= f_1(y_1) + \cdots + f_n(y_n), \end{aligned}$$

where we define

$$f_j(y) := \lambda |y| + \frac{\rho}{2}(v_j - y)^2 \quad \forall\, j = 1, 2, \cdots, n.$$

Analysis of functions f_j

For simplicity of the presentation, we consider the function

$$f(y) = \lambda |y| + \frac{\rho}{2}(v - y)^2.$$

Computing the derivative of $f(y)$ for $y \neq 0$, we have

$$f'(y) = \begin{cases} \lambda - \rho(v - y) & \forall\, y > 0, \\ -\lambda - \rho(v - y) & \forall\, y < 0. \end{cases}$$

Let $f'(y) = 0$. Then we have

$$y = v - \frac{\lambda}{\rho} \text{ for } y > 0 \quad \text{and} \quad y = v + \frac{\lambda}{\rho} \text{ for } y < 0.$$

Therefore, the all critical numbers of f are given by

$$c = v - \frac{\lambda}{\rho} \text{ if } c > 0, \quad c = v + \frac{\lambda}{\rho} \text{ if } c < 0, \quad c = 0.$$

In order to find the minimum of f, we consider the following three cases:

$$v > \frac{\lambda}{\rho}, \quad v < -\frac{\lambda}{\rho}, \quad -\frac{\lambda}{\rho} \leq v \leq \frac{\lambda}{\rho}.$$

Case 1: $v > \frac{\lambda}{\rho}$

In this case, $c = v - \frac{\lambda}{\rho} > 0$ is the only critical number and

$$\begin{aligned} f(c) = f(v - \frac{\lambda}{\rho}) &= \lambda(v - \frac{\lambda}{\rho}) + \frac{\rho}{2}(v - (v - \frac{\lambda}{\rho}))^2 \\ &= \frac{\rho}{2}\left(v^2 - (v - \frac{\lambda}{\rho})^2\right) < \frac{\rho}{2}v^2 = f(0). \end{aligned}$$

For $y \geq 0$, since f is a quadratic polynomial in y with positive leading coefficient, we can conclude that $f(c) \leq f(y)$ for all $y \geq 0$.

For $y < 0$, $f(y)$ is monotone decreasing since

$$\begin{aligned} f'(y) &= \lambda \text{sign}(y) - \rho(v - y) = -\lambda - \rho v + \rho y \\ &< -\lambda - \lambda + \rho y = -2\lambda + \rho y < 0, \end{aligned}$$

which implies $f(y) > f(0)$ for all $y < 0$.

Therefore, f has a minimum at $c = v - \frac{\lambda}{\rho} > 0$.

Case 2: $v < -\frac{\lambda}{\rho}$

In this case, $c = v + \frac{\lambda}{\rho} < 0$ is the only critical number and

$$\begin{aligned} f(c) = f(v + \frac{\lambda}{\rho}) &= -\lambda(v + \frac{\lambda}{\rho}) + \frac{\rho}{2}\left(v - (v + \frac{\lambda}{\rho})\right)^2 \\ &= \frac{\rho}{2}\left(v^2 - (v + \frac{\lambda}{\rho})^2\right) < \frac{\rho}{2}v^2 = f(0). \end{aligned}$$

For $y \le 0$, since f is a quadratic polynomial in y with positive leading coefficient, we can conclude that $f(c) \le f(y)$ for all $y \le 0$.

For $y > 0$, $f(y)$ is monotone increasing since

$$\begin{aligned} f'(y) &= \lambda \text{sign}(y) - \rho(v - y) = \lambda - \rho v + \rho y \\ &> \lambda + \lambda + \rho y = 2\lambda + \rho y > 0, \end{aligned}$$

which implies $f(y) > f(0)$ for all $y > 0$.

Therefore, f has a minimum at $c = v + \frac{\lambda}{\rho} > 0$.

Case 3: $-\frac{\lambda}{\rho} \le v \le \frac{\lambda}{\rho}$

In this case, we have no critical number except the non-differentiable point $y = 0$.

For $y > 0$, we have

$$\begin{aligned} f'(y) &= \lambda - \rho(v-y) = \lambda - \rho v + \rho y \\ &\ge \lambda - \lambda + \rho y = \rho y > 0. \end{aligned}$$

Thus, $f(y)$ is monotone increasing and then $f(y) > f(0)$ for all $y > 0$.

For $y < 0$, we have

$$\begin{aligned} f'(y) &= -\lambda - \rho(v-y) = -\lambda - \rho v + \rho y \\ &\le -\lambda + \lambda + \rho y = \rho y < 0. \end{aligned}$$

Thus, $f(y)$ is monotone decreasing and then $f(y) > f(0)$ for all $y < 0$.

Therefore, f has a minimum at 0.

Solution of problem (P_2)

By the above discussions, we have

$$\arg\min_{y} f(y) = \begin{cases} v + \frac{\lambda}{\rho}, & \text{if } v < -\frac{\lambda}{\rho}, & \text{(case 2)} \\ 0, & \text{if } |v| \le \frac{\lambda}{\rho}, & \text{(case 3)} \\ v - \frac{\lambda}{\rho}, & \text{if } v > \frac{\lambda}{\rho}. & \text{(case 1)} \end{cases}$$

In other words, we have

$$\arg\min_{y} f(y) = \mathcal{S}_{\lambda/\rho}(v) = \text{sign}(v) \max(0, |v| - \lambda/\rho).$$

Therefore,

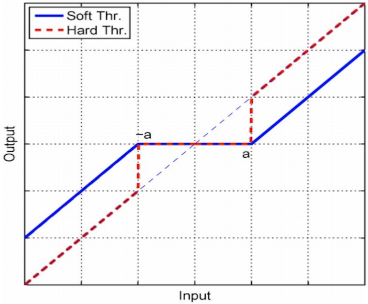
$$\boldsymbol{y}^{(i+1)} = \arg\min_{\boldsymbol{y}} F_2(\boldsymbol{y}) = \mathcal{S}_{\lambda/\rho}(\boldsymbol{v}) = \mathcal{S}_{\lambda/\rho}(\boldsymbol{z}^{(i+1)} + \boldsymbol{u}^{(i)}).$$

where *the soft-thresholding function*,

$$\mathcal{S}_{\lambda/\rho}(\boldsymbol{v}) := \text{sign}(\boldsymbol{v}) \odot \max(\boldsymbol{0}, \ |\boldsymbol{v}| - \lambda/\rho),$$

and $\text{sign}(\cdot)$, $\max(\cdot, \cdot)$, and $|\cdot|$ are all applied to the input vector $\boldsymbol{v}$ component-wisely, and $\odot$ is the Hadamard product.

Soft- and hard-thresholding functions



(This figure quoted from "M. Elad, M. A. T. Figueiredo, and Y. Ma, On the role of sparse and redundant representations in image processing, *Proceedings of the IEEE*, 98 (2010), pp. 972-982")

- soft-thresholding function: continuous
- hard-thresholding function: discontinuous