$$\min_{x \in \mathbb{R}^n} \|x\|_1 + \frac{1}{2\alpha}\|x - y\|_2^2$$

where $\alpha > 0$ is a given scalar
y - is a known vector

Task 5
HW 4
Optimization

Page 1

1. Analyze the convexity

1.1. Feasible set is convex: $x \in \mathbb{R}^n$

1.2. $\|x\|_1$ - convex

$\|x - y\|_2^2$ - convex

and we know that the sum of convex functions with non-negative coefficiens is convex function

$\underbrace{1}_{>0} \cdot \|x\|_1 + \boxed{\frac{1}{2\alpha}} > 0$ because $\alpha > 0$ ~~according~~ by condition

therefore objective function is convex

1.3 If feasible set is convex ~~and~~ and objetive function is convex
=> we have convex optimization problem.

2. Analyze smothness

2.1 $\|x\|_1 = \sum_i |x_i|$ - not smoth function, not differentiable

2.2. $\|x - y\|_2^2$ - smoth function

$\Downarrow$

2.3. Objective function $\|x\|_1 + \frac{1}{2\alpha}\|x - y\|_2^2$ -> not smoth (not differentiable) function

Task 5
HW4
Optimization

Page 2

3. Compose equivalent differentiable problem

$$\min_{x \in \mathbb{R}^n} \|x\|_1 + \frac{1}{2\alpha} \|x - y\|_2^2$$

because $\|x\|_1 = \sum_i |x_i|$

we can introduce an additional variable t

$$|x_i| = t_i$$

A) $\min_x \|x\|_1 + \frac{1}{2\alpha} \|x - y\|_2^2$

$$\Downarrow$$

$$\min_{(x,t)} \sum_i t_i + \frac{1}{2\alpha} \|x - y\|_2^2$$

s.t. $|x_i| = t$

B) We can ~~in~~ replace equalities on inequalities

$|x_i| = t_i \Rightarrow$

$|x_i| \leq t_i$

and we get an equivalent optimization problem with constraints with inequality.

$$\min_{(x,t)} \sum_i t_i + \frac{1}{2\alpha} \|x - y\|_2^2$$

s.t $|x_i| \leq t_i$ $\begin{cases} x_i \leq t_i \\ -x_i \leq t_i \end{cases}$

in a differentiable form.

Task 5
HW4
Optimization

Page 3

4. Write the KKT conditions:

$$\min_{(x,t)} \frac{1}{2\alpha} \|x-y\|_2^2 + \sum_i t_i$$

$$\text{s.t} \quad |x_i| \leq t_i \begin{cases} x_i \leq t_i \\ -x_i \leq t_i \end{cases}$$

$$x_i - t_i \leq 0$$

$$-x_i - t_i \leq 0$$

$$\mu_i (x_i - t_i) = 0$$

$$\nu_i (-x_i - t_i) = 0$$

$$\mu_i, \nu_i \geq 0$$

$$\nabla \mathcal{L}_{t,x} (x, t, \mu, \nu) = 0$$

where

$$\mathcal{L}(x, t, \mu, \nu) = \frac{1}{2\alpha} \|x-y\|_2^2 + \sum_i t_i + \mu_i \sum_i (x_i - t_i) + \nu_i \sum_i (-x_i - t_i)$$

5. Solve the KKT system and reconstruct the solution of the original problem. Pay special to the dependence of x^* on α and y.

Task 5
HW 4
Optimization

$$L = \frac{1}{2\alpha}\|x-y\|_2^2 + \sum_i t_i + \sum_i \mu_i (x_i - t_i) + \sum_i \nu_i (-x_i - t_i)$$

$$\|x-y\|_2^2 = \sum_i (x_i - y_i)^2$$

$$\frac{\partial L}{\partial x_k} = \frac{1}{2\alpha} 2 (x_i - y_i) + \mu_k - \nu_k = 0 \qquad (x_k - y_k) + \alpha\mu_k - \alpha\nu_k = 0$$

$$\frac{\partial L}{\partial t_k} = 1 - \mu_k - \nu_k = 0 \qquad \mu_k + \nu_k = 1$$

if $\mu_k = 0$ $\nu_k = 1$

if $\nu_k = 0$ $\mu_k = 1$

I

$\mu_i = 0$ $\nu_i = 1$

$x_i - y_i - \alpha\nu_i = 0$

$x_i - y_i = \alpha$ $x_i = y_i + \alpha$

$\nu_i(-x_i - t_i) = 0$

$\nu_i = 1$: $-x_i - t_i = 0$

$x_i = -t_i$

$$\boxed{\begin{aligned} \mu_i &= 0 \\ \nu_i &= 1 \\ x_i &= -t_i \\ x_i &= y_i + \alpha \\ t_i &= -y_i - \alpha \end{aligned}}$$

$\mu_i, \nu_i \geq 0$ ok

$x_i - y_i - \alpha = 0$ ok

$1 - 0 - 1 = 0$ ok

II

$\mu_i = 1$ $\nu_i = 0$

$x_i - y_i + \alpha \cdot 1 = 0$

$x_i = y_i - \alpha$

$\mu_i(x_i - t_i) = 0$

$\mu_i = 1$: $x_i = t_i$

$$\boxed{\begin{aligned} \mu_i &= 1 \\ \nu_i &= 0 \\ x_i &= t_i \\ x_i &= y_i - \alpha \\ t_i &= y_i - \alpha \end{aligned}}$$

$\mu_i, \nu_i \geq 0$ ok

$x_i - y_i + \alpha = 0$ ok

$1 - 1 - 0 = 0$ ok.

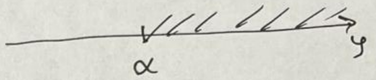
5 Part II reconstruct the solution of the original problem

Task 5
HW4
Optimization
Page 5

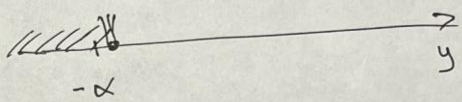
Because in the original problem we have $\|X\|_1 = \sum_i |X_i|$ and since a module is a "special" function, we need to consider under what conditions $X=0$, $X>0$ and $X<0$.

Now we know, that (I) $X_i = y_i + \alpha$ $(\alpha > 0)$
(II) $X_i = y_i - \alpha$

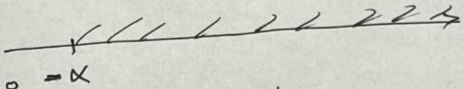
① If $y_i > \alpha \Leftrightarrow X_i = y_i - \alpha > 0$
$y_i - \alpha > 0$



② if $y_i < -\alpha \Leftrightarrow X_i = y_i + \alpha < 0$
$y_i + \alpha < 0$

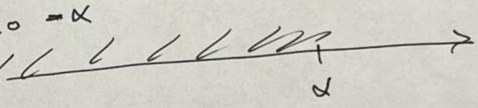


③ if $y_i \geq -\alpha$

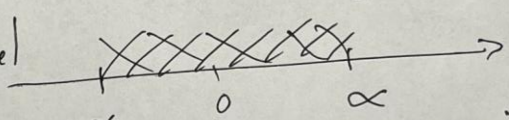


$X_i = (y_i + \alpha) > 0$

$y_i \leq \alpha$ $(y_i - \alpha) \leq X_i < 0$



then consider the interval $-\alpha < y < \alpha$



in which we have contradiction $X_i < 0$ and $X_i > 0$ simultaneously.
To eliminate the contradiction, it is necessary to define X equal to zero for $-\alpha \leq y_i \leq \alpha$.

$$X_i^* = \begin{cases} y_i - \alpha, & y_i > \alpha \quad (X > 0) \\ 0, & |y_i| \leq \alpha \quad (-\alpha < y < \alpha) \\ y_i + \alpha, & y_i < -\alpha \quad (X < 0) \end{cases}$$

6. Explain why this point is a solution or provide additional analysis

Task 5
HW4
Optimization

Consider the behavior of the objective function on a feasible set. Since our function is convex on a feasible set, we need to consider three intervals:

1. At $x<0$, our convex function decreases monotonically (gradient <0, Hessian is positive definite everywhere). For a convex function on the interval of monotonous decrease - any critical point - a minimum point. (When the function is limited only to the right). In our case, such a point is $x = y+\alpha$

2. When $x=0$. As we see from point 5 of this solution, y_i takes values from $-\alpha$ to $+\alpha$, and for $x=0$, the value $f(0)=\frac{y^2}{2\alpha}$. If we consider the y as a function of α (quadratic in α), we get that the value of our objective function is proportional to the value of the constant α - this means that the function itself is constant at this interval, with a predermined α, as provided in the condition of the problem. That is, the function is constant (for each α) and $x=0$ is a solution minimizing our objective function.

3. For $x>0$, our convex function increases monotonically (gradient >0, Hessian is positive everywhere). For a convex function on a interval of monotonic increase, any critical point is a minimum point (when the function is limited to the left and not limited to the right). In our case, this is the point $x=y-\alpha$.

Task 5
HW4
Optimization

Page 6.2

In addition to what was written in part 6 on page 6, I would like to note that in KKT we have defferetiable convex problem with affine constraints.
And this means that we have the necessary and sufficient conditions for the found point x^* to be the solution of the optimal problem.

7. Plot the solution in the case of $n=1$ and explain why the problem is called "soft thresholding".

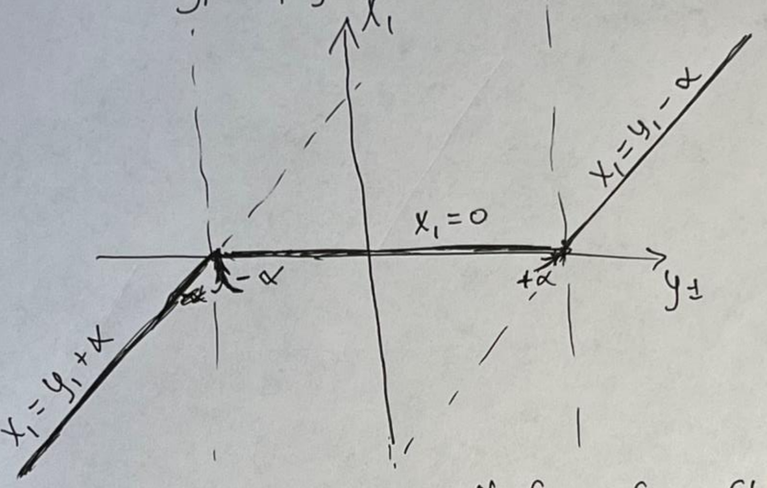
Bonus

Task 5
HW4
Optimization

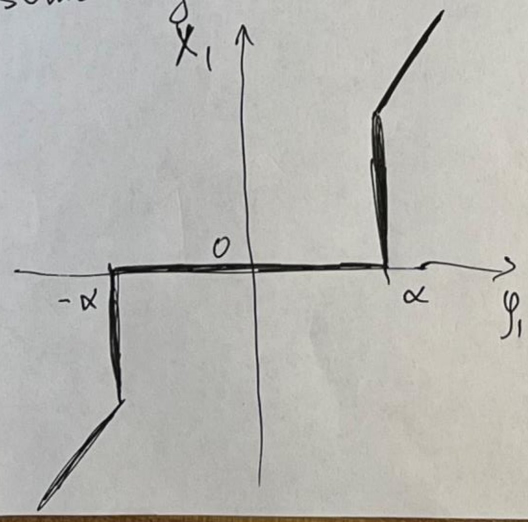
Page 7

$n=1$

$$x_1^* = \begin{cases} y_1+\alpha, & y_1<-\alpha \quad (x<0) \\ 0, & |y_1|\leqslant\alpha \quad (-\alpha<y<\alpha) \\ y_1-\alpha, & y_1>\alpha \quad (x>0) \end{cases}$$



As we can see, the solution has the form of a soft threshold, in contrast to the case of a hard threshold, which looks something like this:



And since the graphical representation of the solution to this problem is in the form of a soft threshold, the problem is called a "soft threshold" problem.