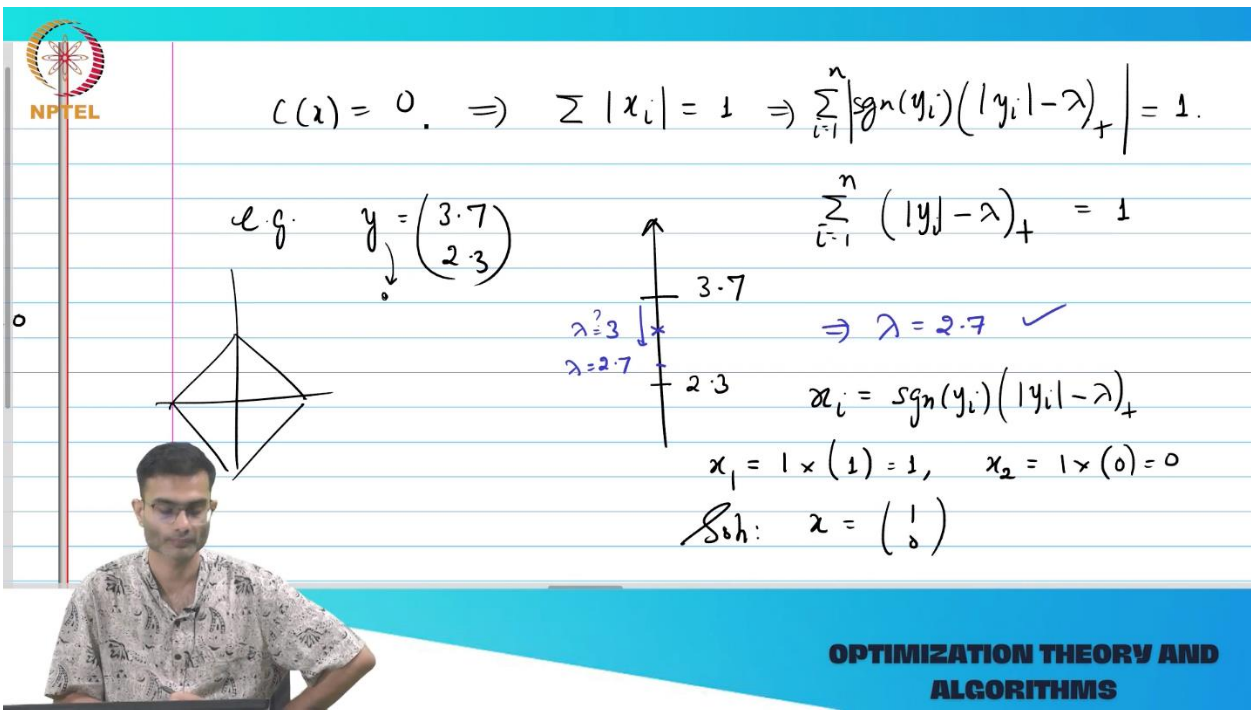
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 11
Lecture - 76

Soft thresholding example

Down to expression b, I still have not got the value of x_i that I wanted, how do I get that? We have not yet used the fact that $C(x)$ has to be equal to 0, right? So, let us use that; it is the only missing part of the puzzle. $C(x) = 0$ implies that $|x_i|$ should sum to 1, which implies that I have to substitute what was in the previous expression, right? So, signum(y_i) and $y_i - \lambda$ plus this should sum to 0, right? $\sum_{i=1}^{n} |y_i - \lambda| = 0$.

In principle, does this look like an expression that I can solve? Look at each one of those terms. So, I will give you y, it will be an n-dimensional vector, which means there are n numbers inside it. Take each one of them; take y_1, do you know sign(y_1)? Yes, it is given to you, it will be $+1$ or -1, right? Then $y_i - \lambda$. So, you can calculate what that is with the plus operator.

Finally, what will you get when I add n of these terms? Will I be able to solve and calculate λ? Yes, it is just a linear equation and in each term λ will come or may not come depending on the plus operator, and you solve it to calculate λ. There is no closed-form expression, but it is very easy to solve, right? So, the unusual thing is there is no closed-form expression, but it is incredibly easy to solve.



So, we will take an example to clarify this. The signum is appearing here. The signum function is capturing the cases of signum(y_i), because the signum is not being applied to x_i, it is applied to

y_i. $|x_i|$ is being rewritten as this. No, oh sorry, $|\cdot|$ of the whole thing, yeah, you are right, $|\cdot| = 1$. So, we will simplify this, okay?

Let us take an example to make this more concrete. So, let us take some random values. So, this is my point, this is my constraint, right? This is my right $(2.3, 3.7)$, which is some point over here, and now I want to find out what the projection of this is on the L_1-ball. That is the problem, okay? So, there are two terms in this summation, right?

Now, all of the y_i's are positive, so the signum becomes just 1, and I am taking the modulus of it. So, actually, this signum function will go away because there is a modulus around it. This will simply become:

$$\sum_{i=1}^{n} |y_i - \lambda| + \text{signum}(y_i) = 0.$$

When $y_i = 0$, the signum makes it 0. No, no, you won't need that. The plus operator takes care of it. Okay, now this is the expression that I have to work with. Let us plug in y_1 and y_2 one by one, right?

So, what is it? I can do it like that, but then I don't know how to evaluate the plus operator because I don't know λ. I put 3.7 and 2.3, I have $3.7 - \lambda$, plus the operator. I do not know how to evaluate that because I do not know λ. So, there is a very simple graphical way of doing this again. Let us take a vertical scale, let us mark 3.7 over here, and let us mark 2.3 over here.

Now, each of these numbers, so you need to pay attention, each of these numbers I am subtracting λ from it, and then applying the plus operator to it. That thing should sum to 1. So, supposing I take my λ over here, okay?

Let us say that this is $\lambda = 3$, question mark. So, the first term, $3.7 - 3$, what is the plus operator going to spit out? 0.7. The next term, what will it spit out? It does not sum to 1, right? So, now, imagine that I have a slider available to me and I need to make λ greater or smaller. I need to lower it because the sum is not yet coming to 1.

So, supposing I lower it all the way to 2.7, right? If I bring this down to, for example, over here, $\lambda = 2.7$, then what will I get? I will get:

$$3.7 - 2.7 = 1, \quad 2.3 - 2.7 = 0.$$

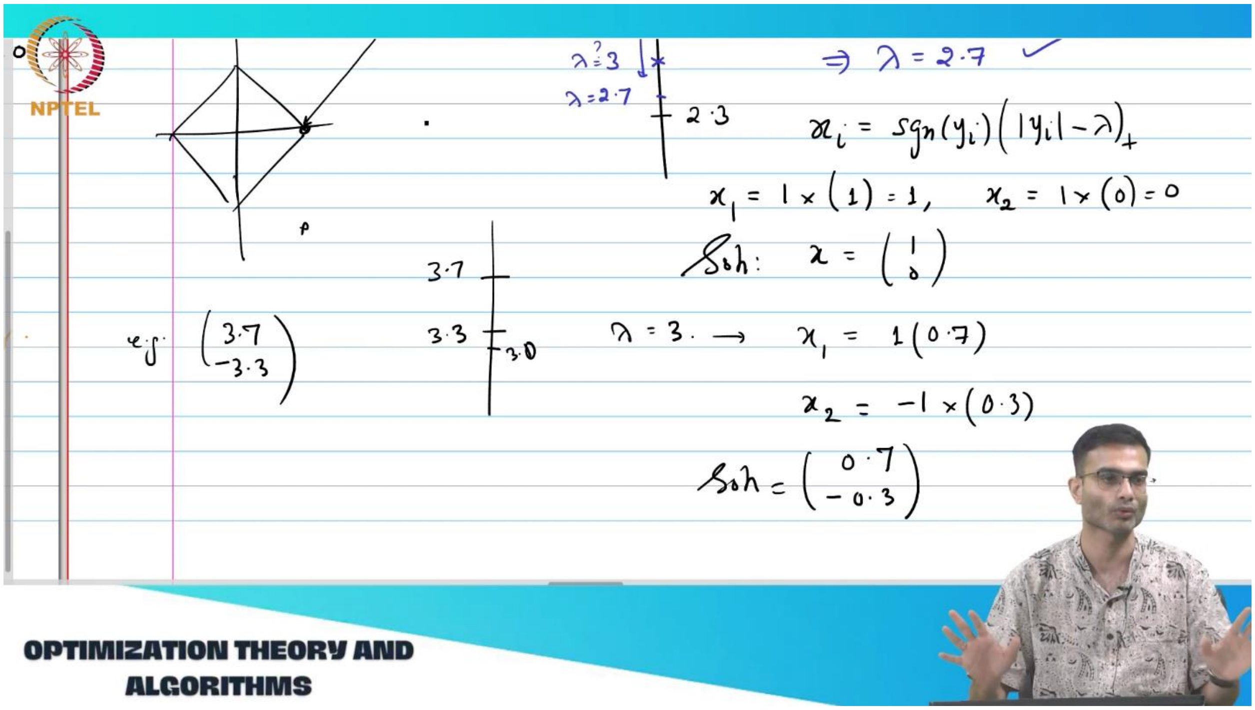
Done. If I go any further than 2.7, what happens? The first term will become greater than 1, and the second term will continue to be 0, right?

So, there is a sweet spot over here. This implies that $\lambda = 2.7$ is the correct solution. So, I have got λ right. Finally, I need to get x, right? λ is the intermediate variable. So, I had that expression for x_i, which was:

$$x_i = \text{signum}(y_i)|y_i - \lambda|.$$

So, what will x_1 be? $\text{signum}(3.7) = 1$, and $|3.7 - 2.7| = 1$. For x_2, $\text{signum}(2.3)$ is 1, and $|2.3 - 2.7| = 0$.

So, $x = 1,0$. The solution is $x = (1,0)$, right? So, actually, this projects to this point over here. Right? So, you got it. We will take one more example, this time with a negative coordinate to further drill this in. So, let us take an example of, let us say, $(3.7, -3.3)$. This is the point that I am trying to project. So, which quadrant is this? Bottom somewhere here, and I am trying to project this into the L_1-ball, right?



We will quickly wrap up from here. So, first step, right, we draw this over here. Do I need to draw? See, remember there is a modulus around y_i. So, on the scale, what should I draw? The modulus of the y_i-values.

So, the modulus of the y_i-values is going to give me 3.7 and 3.3. See, there is a modulus happening over here: $|y_i|$. So, each y_i can be negative, but it does not matter; it is $|y_i - \lambda|$ over here, right?

So, what value of λ do you think will work? It has to be less than 3.3. Supposing I stop at, let us say, $\lambda = 3.1$, will it work?

$$3.7 - 3.1 = 0.6, \quad 3.1 - 3.3 = -0.2.$$

It still does not sum to 1. So, you can figure out what is the best value of λ, then? $\lambda = 3$, right?

So, $\lambda = 3$ is what works, and then you apply the method. You can go through this exercise, and you will get the solution:

$$x_1 = \text{signum}(3.7) \times |3.7 - 3| = 1 \times 0.7 = 0.7,$$

$$x_2 = \text{signum}(-3.3) \times |-3.3 - 3| = -1 \times 0.3 = -0.3.$$

So, the solution is $x = (0.7, -0.3)$.

This is simple enough that you could actually write a clever enough piece of code that does it.

Perpendicularity does not make sense if the tangent direction is not defined. Perpendicular to what? To any of, like that. Okay, so this we had to rush a little bit in the end because, I mean, this topic of subgradients and projection on the L_1-ball is now over. But this is a new concept for most of you, I think. Most of you are seeing the idea of subgradients for the first time, and this is particularly useful if you go more into numerical analysis. It is a very, very important tool. So, just go over it once. It is not too tough once you get familiar with it, okay?