NOM BOUKHELOUA

Prénom Rémi

Promo 2020

Date

20150257: BOUKHELOUA Rémi
M1:
ST2CXO-DE (18/04/2019)
Amphi orange

14,75/20

MATIÈRE Optimisation (1/2)

Ex 5.

We have, in x_0, $f(x) \approx f(x_0) + f'(x_0)(x - x_0)$. ⚠

As we look for minimum we look for $f(x) = 0$.

So $f(x_0) + f'(x_0)(x - x_0) = 0$.

$$x = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Recursively : $x_{k+1} = x_k - \dfrac{f(x_k)}{f'(x_k)}$ ok

We want to optimise a function f, so we look for $\nabla f = 0$.

(and $H_f(x^{(k)}) > 0$ if f not convex).

Thus, we must apply Newton's method on $\nabla f \overset{\text{déf}}{=} F(x)$

So we have : $x_{k+1} = x_k - \dfrac{\nabla f(x_k)}{\nabla f'(x_k)}$ Not exactly (not the gradient)

But $f' = \nabla f$, so we have $x_{k+1} = x_k - \dfrac{\nabla f(x_k)}{\nabla^2 f(x_k)}$.

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

That's what we call the Hessian matrix of f H_f

So we have : $x_{k+1} = x_k - \dfrac{\nabla f(x^{(k)})}{H_f(x^{(k)})}$.

$\frac{a}{b} = a \times b^{-1}$. And as we have to care about order for matrix multiplication :

We obtain $x_{k+1} = x_k - [H_f(x^{(k)})]^{-1} \nabla f(x^{(k)})$

2

Ex 6. If we have a constrained optimization problem such as

~~min ||x||~~ $\min_{x \in \Omega} f(x)$

subject to $g_i(x) \leq 0$ with $i \in \mathbb{R}$.

we could decide to switch to an unconstrained optimization method introducing a penalty $P(x)$ such as

cost (or objective) function to optimize is: $h(x) = f(x) + \gamma P(x)$.

where γ is a penalty rate strictly positive.

and $P(x) = \begin{cases} 0 & \text{if } x \in \Omega \\ \sum \max(0, g_i) & \text{else} \end{cases}$

not necessarily, but why n

In that case we would not use constrains but if a value is not respecting ~~g conditions~~ $g_i(x) \leq 0$ condition, ~~we~~ it would be that impacted that we ~~would~~ ~~wear~~ ~~p~~ are sure it exists a ~~~~ lower value of $h(x)$ where x is respecting ~~condit~~ constraints.

OK

2

Ex 1

1. $f_1(x) = \sin(x)\cos(x)$, $x \in \mathbb{R}$.

$f_2(x, y) = 18x - 7y$, $(x, y) \in \mathbb{R}^2$

Let ~~us consider~~ us consider a and b such as $a, b \in \mathbb{R}^2$.

According to midpoint inequality; f is convex if

$$f\left(\frac{a+b}{2}\right) \leq \frac{f(a)}{2} + \frac{f(b)}{2}.$$

Only if f is continuous (it should be noted)

$f_2\left(\frac{a_x+b_x}{2} ; \frac{a_y+b_y}{2}\right) = 18\frac{(a_x+b_x)}{2} - 7\frac{(a_y+b_y)}{2} \quad = 9(a_x+b_x) - \frac{7}{2}(a_y+b_y)$

$\frac{1}{2} f_2(a) = 9a_x - \frac{7}{2}a_y$ and $\frac{1}{2} f_2(b) = 9b_x - \frac{7}{2}b_y$ ⚠ Calculus error

$f_2\left(\frac{a_x+b_x}{2} ; \frac{a_y+b_y}{2}\right) - \left(\frac{1}{2} f_2(a) + \frac{1}{2} f_2(b)\right) = \cancel{18}^{\,9}(a_x+b_x) - \cancel{7}^{\,\frac{7}{2}}(a_y+b_y) - \left[9a_x - \frac{7}{2}a_y + 9b_x - \frac{7}{2}b_y\right]$

$= \cancel{9a_x - \frac{7}{2}a_y + 9b_x - \frac{7}{2}b_y} = 0$

0,25

~~We can't~~ As $a, b \in \mathbb{R}^2$, we can't conclude on the sign of $f_2\left(\frac{a+b}{2}\right) - \frac{1}{2}f_2(a) - \frac{1}{2}f_2(b)$, so f_2 is neither convex or concave.

f_2 is convex and concave

$f_3(x) = \sqrt{x}$, $x \in \mathbb{R}_{>0}$

f_3 is continuous and ~~differentiable~~ twice differentiable on $\mathbb{R}_{>0}$

so we can say that f_3 is convex if $f_3''(x) \geq 0$

$f_3'(x) = \frac{1}{2} \times 1 \times \frac{1}{\sqrt{x}} = \frac{1}{2\sqrt{x}}$

$f_3''(x) = -1 \times \cancel{2} \times \frac{1}{\cancel{2}\sqrt{x}} \times \frac{1}{4x}$

$= \frac{-1}{4x\sqrt{x}}$

1

As $x \in \mathbb{R}_{>0}$ $\quad 4x\sqrt{x} > 0$ and $f_3''(x) \leq 0$.

$-f_3''(x) \geq 0$.

~~So~~ The opposite of f_3 is convexe, so f_3 is concave

$f_4(x,y) = 1 - \left(\frac{x^2+y^2}{2}\right)$, $(x,y) \in \mathbb{R}^2$

Let us consider a and b such as $a, b \in \mathbb{R}^2$ (*)

f_4 is convex if $f_4\left(\frac{a+b}{2}\right) \leq \frac{f_4(a)}{2} + \frac{f_4(b)}{2}$. ⚠ Only if f is continuous. (here, it is, but it should be noted)

$f_4\left(\frac{a+b}{2}\right) = 1 - \frac{\left(\frac{a_x+b_x}{2}\right)^2 + \left(\frac{a_y+b_y}{2}\right)^2}{2}$

$\frac{1}{2} f_4(a) = \frac{1}{2}\left(1 - \frac{a_x^2 + a_y^2}{2}\right)$ and $\frac{1}{2} f_4(b) = \frac{1}{2}\left(1 - \frac{b_x^2+b_y^2}{2}\right)$

$f_4\left(\frac{a+b}{2}\right) - \frac{1}{2}f(a) - \frac{1}{2}f(b) = \cancel{1} - \frac{\left(\frac{a_x+b_x}{2}\right)^2 + \left(\frac{a_y+b_y}{2}\right)^2}{2} - \cancel{\frac{1}{2}}\left[\cancel{1} - \frac{a_x^2+a_y^2+b_x^2+b_y^2}{4}\right]$

$= \frac{-(\cancel{a_x^2} + \cancel{b_x^2} + 2a_xb_x + \cancel{a_y^2} + \cancel{b_y^2} + 2a_yb_y) + 2a_x^2 + 2a_y^2 + 2b_x^2 + 2b_y^2}{8}$

$= \frac{a_x^2 - 2a_xb_x + b_x^2 + a_y^2 - 2a_yb_y + b_y^2}{8}$

$$f_u\left(\frac{a+b}{2}\right) - \frac{f(a)+f(b)}{2} = \frac{(a_x - b_x)^2 + (a_y - b_y)^2}{8}$$

0,75

For every $a, b \in \mathbb{R}^2$, $(a_x - b_x)^2 \geq 0$ and $(a_y - b_y)^2 \geq 0$

so $f_u\left(\frac{a+b}{2}\right) - \frac{f(a)+f(b)}{2} \geq 0$.

$$f_u\left(\frac{a+b}{2}\right) \geq \frac{f(a)}{2} + \frac{f(b)}{2}.$$

So the line is under or on the graph what means

f_u is concave.

Only if f is continuous

(IT HAS BEEN NOTED ⊛)

2. $\{x \in \mathbb{R}^n : |x - c| \leq r\}$.

Let us consider $x, y \in \mathbb{R}^n$ such as they are in or on the closed ball, and $\lambda \in [0;1]$.

The closed ball is a convex set if $\lambda x + (1-\lambda)y$ is in the set.

$|x - c| \leq r \iff \lambda|x - c| \leq r \times \lambda$

As $\lambda \geq 0$ $|x - c| \leq r \iff |\lambda x - \lambda c| \leq \lambda r$.

$|y - c| \leq r \iff (1-\lambda)|y - c| \leq (1-\lambda)r$

As $\lambda \leq 1$, $1-\lambda \leq 0$, $|y - c| \leq r \iff |(1-\lambda)y - (1-\lambda)c| \leq (1-\lambda)r$

We add both inequalities and call it I

$I \iff |\lambda x - \lambda c| + |(1-\lambda)y - (1-\lambda)c| \leq \lambda r - (1-\lambda)r$.

$|a| + |b| \geq |a+b|$ (Triangular inequality) yes

so $|\lambda x - \lambda c + (1-\lambda)y - (1-\lambda)c| \leq |\lambda x - \lambda c| + |(1-\lambda)y - (1-\lambda)c|$

$\leq r$.

$|\lambda x + (1-\lambda)y - c| \leq r$.

So $\lambda x + (1-\lambda)y$ satisfies equation of a point in or on the closed ball for $\lambda \in [0;1]$.

We can conclude that the closed ball is a convex set.

1

NOM BOUKHELOUA

Prénom Rémi

Promo 2020

Date

MATIÈRE Optimisation (2/2)

top and bottom

Ex 2

1. ~~We are~~ Optimization variables : r and h.

~~Co~~ Cost function : $\min_{(r,h)\in\mathbb{R}^2}$ $\cancel{\pi r^2 h}$ $2\pi r h$ $+ 2\pi r^2$

1

subject to $r \geq 0$, $h \geq 0$.

$\pi r^2 \times h = \cancel{1} 33cL = 100\,cm^3$

$\cancel{\pi r^2 \times h = 100\,cm^3}$.

2. Optimization variables: a, b, c where a is number of type A drink produced, b Type B drink and c Type C drink.

Objective function : $\max_{(a,b,c)\in\mathbb{N}^3}$ $0{,}55a + 0{,}65b + 0{,}85c$.

subject to : $3a + 2b + 2c \leq 1700$.

$2a + 3c \leq 900$

$4b + 2c \leq 1350$.

$a \geq 0$, $b \geq 0$, $c \geq 0$.

1,5

Ex 4 : 1. $x^{k+1} = x^k - \alpha \nabla f(x^k)$ ~~w~~

$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3} \right)$

$= \cancel{2x_1 - 12 ; 4x_3 + 618x_2^3}$

$\nabla f = (2x_1 - 12 ; 8x_2^3 + 72x_2^2 + 160x_2 + 216 ; 12x_3^3 - 108x_3^2 + 324x_3 - 324)$

no need to expand

$x^1 = x^0 - \alpha \nabla f(x^0)$

$= (10 \quad -2 \quad 3) - \frac{1}{4}(8 ; \cancel{120} ; \cancel{1944})$ 8 12

$= (10 \quad -2 \quad 3) - (2 \quad \cancel{30} \quad 486)$

$= (8 \quad \cancel{-32} \quad -483)$ $(8 \quad -4 \quad 3)$

The calculus would have been easier if you had kept the factorised

1

2. In the steepest descent method, we don't have a ~~k~~ fixed learning rate α. At each iteration we recalculate as following : $\alpha_{k} = \underset{\alpha \geq 0}{\operatorname{argmin}}\ f(x^{k} - \alpha \nabla f(x^{k}))$.

So at the beginning, it's a bigger one and it ~~do~~ become lower and lower during algorithm. That protect from a big one jumping over minimum over and over.

It depen

0,75

3. In stochastic variant we ~~calculate a function f~~ ~~for~~ divide dataset into partitions 0 to T. and we calculate a function f_i for each partitions. In single batch we go over each partition but in minibatch we select randomly one or many partitions

(and estimate f with corresponding f_i

1,5 It improve greatly the time gradient descent takes in a big dataset case.

OK

Ex3.

2. a. Canonical form:

Optimization variables: x_1, x_2, x_3

Objective function ~~obj~~ $\max_{(x_1,x_2,x_3)\in\mathbb{R}^3} \; 5x_1 + 3x_2 + 9x_3$.

$$\text{subject to} \quad \begin{aligned} 6x_1 - 6x_2 + 2x_3 &\leq 20 \\ x_1 + 9x_2 + 4x_3 &\leq 60 \\ -2x_1 + 2x_2 + 4x_3 &\leq 40. \\ x_1 \geq 0,\ x_2 \geq 0,\ &x_3 \geq 0. \end{aligned}$$

Should be in matricial form.

Standard form:

forms

$$\max_{(x_1,x_2,x_3)\in\mathbb{R}^3} \; 5x_1 + 3x_2 + 9x_3.$$

$$\text{subject to} \quad \begin{aligned} 6x_1 - 6x_2 + 2x_3 + x_4 &= 20 \\ x_1 + 9x_2 + 4x_3 + x_5 &= 60 \\ -2x_1 + 2x_2 + 4x_3 + x_6 &= 40. \end{aligned}$$

$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0, x_6 \geq 0.$

0,5

b.

b.	x_1	x_2	x_3	x_4	x_5	x_6	b.	
→ x_4	6	-6	2	1	0	0	20	10
x_5	1	9	4	0	1	0	60	15
x_6	-2	2	4	0	0	1	40	10.
Z.	5	3	9	0	0	0	0.	

↑ bigger one.

x_3	3	-3	1	$\frac{1}{2}$	0	0	10

...nds...

More or less...

	x_1	x_2	x_3	x_4	x_5	x_6	b	
x_3	3	-3	1	$\frac{1}{2}$	0	0	10	$-\frac{10}{3}$
$\cancel{x_5}$	-11	21	0	$\cancel{-\frac{1}{2}}$ -2	$\cancel{0}$ 1	0	20	$\frac{20}{21}$
x_6	-14	14	0	-2	0	1	0	0
Z	-22	30	0	-2	0	0	-90	

bigger

Error in the selection of variable!

x_2 -1 1 0 $-\frac{1}{7}$ 0 $\frac{1}{14}$ 0

	x_1	x_2	x_3	x_4	x_5	x_6	b
x_3	0	0	1	$\frac{1}{14}$	0	$\frac{3}{14}$	10
x_5		0	0				20
x_2	-1	1	0	$-\frac{1}{7}$	0	$\frac{1}{14}$	0
Z							-90

1,5

x_2 $-\frac{11}{21}$ 1 0 $-\frac{2}{21}$ $\frac{1}{21}$ 0 $\frac{20}{21}$

	x_1	x_2	x_3	x_4	x_5	x_6	b
x_3	$\frac{10}{7}$	0	1	$\frac{3}{14}$	$\frac{1}{7}$	0	$\frac{230}{21}$
x_2	$-\frac{11}{21}$	1	0	$-\frac{2}{21}$	$\frac{1}{21}$	0	$\frac{20}{21}$
x_6	$-\frac{305}{21}$	0	0				
Z		0	0				