Numerical optimization with python - Dry HW 1

Roni Ben-Dom 207576463

Part 1 - Modeling optimization problems:

1. Denote unknown variables as x_1, x_2, x_3 where:

x_1 - # of kg of vanilla ice cream

x_2 - # of kg of choclate ice cream

x_3 - # of kg of strawberry ice cream

The total cost to maximize: $7x_1 + 6x_2 + 5x_3$

constraint of machine production: $x_1 + x_2 + x_3 \leq 100$

constraint on milk supplies: $0.5x_1 + 0.2x_2 + 0.4x_3 \leq 45$

constraint on cream supplies: $0.4x_1 + 0.7x_2 + 0.4x_3 \leq 60$

constraint on vanilla supplies: $0.1x_1 \leq 10$

constraint on choclate supplies: $0.1x_2 \leq 10$

constraint on strawberries supplies: $0.2x_3 \leq 15$

non negativity constraint: $x_i \geq 0 \quad \forall i \in \{1,2,3\}$

Therefore, we arrived at the following formulation:

$\max(7x_1 + 6x_2 + 5x_3)$ subject to:

- $x_1 + x_2 + x_3 \leq 100$
- $0.5x_1 + 0.2x_2 + 0.4x_3 \leq 45$
- $0.4x_1 + 0.7x_2 + 0.4x_3 \leq 60$
- $0.1x_1 \leq 10 \Rightarrow x_1 \leq 100$
- $0.1x_2 \leq 10 \Rightarrow x_2 \leq 100$
- $0.2x_3 \leq 15 \Rightarrow x_3 \leq 75$
- $x_i \geq 0 \quad \forall i \in \{1,2,3\}$

2. Denote unknown varibles a p_i, $i \in \{1,2,3,4\}$

The total value to maximize: $-\sum_{i=1}^{4} p_i \log(p_i)$

non-negativity constraint: $\forall i \in \{1,2,3,4\}: p_i \geq 0$

probability distribution constraint: $\sum_{i=1}^{4} p_i = 1$

Expected value constraint: $\sum_{i=1}^{4} p_i x_i = 0$

Mathematical formulation:

$\max\left(-\sum_{i=1}^{4} p_i \log(p_i)\right)$ subject to:

- $p_i \geq 0 \quad \forall i \in \{1,2,3,4\}$
- $\sum_{i=1}^{4} p_i = 1$
- $-10.2\,p_1 + 0.4\,p_2 + 16.6\,p_3 + 10.3\,p_4 = 0$ (in general case: $\sum_{i=1}^{4} p_i x_i = 0$)

3. Let $\underbrace{x = (x_1, \dots x_n)}_{\text{unknown variable}}$. We want to find x that minimizes $\|x - p\|_2$ for a given $p \in \mathbb{R}^n$; x ~~[illegible]~~ is in then northen half closed unit ball.

Mathematical formulation:

$\min(\|x - p\|_2)$ subject to:

$\sqrt[2]{x_1^2 + x_2^2 + \dots + x_n^2} \leq 1$

$x_i \geq 0$ for all $i \in \{1, \dots n\}$

a. fullfiled in the last step.

b. now we want $\|x - p\|_1$ to be minimal:

$\min(\|x - p\|_1)$ subject to:

$|x_1| + |x_2| + \dots + |x_n| \leq 1$

$x_i \geq 0$ for all $i \in \{1, \dots, n\}$

~~The definition of the problem [illegible] with the norm, but, given the second constraint we can look at the first one as [illegible] and then the problem is smooth~~

The problem is not smooth due to the absolute value.

3.C. now let $x=(x_1, x_2, \dots, x_n, x_{n+1}, \dots, x_{2n})$ be the unknown variable. $P=(P_1, \dots, P_n, P_1, \dots, P_n)$

we want to minimize $\|x-p\|_1$ and have a smooth problem with the same interpretation for $x_1, \dots, x_n$.

Mathematical formulation:

$\min(\|x-P\|_1)$ subject to:

$|x_1|+|x_2|+\dots+|x_n| \leq 1$

$x_i \geq 0 \quad \forall i \in \{1, \dots, n\}$

$x_i = x_{i+n} \quad \forall i \in \{1, \dots, n\}$

The third constraint creates the next situation:

$$|x_1-P_1|+|x_2-P_2|+\dots+|x_n-P_n|+|\underset{\overset{\shortparallel}{x_1}}{x_{n+1}}-\underset{\overset{\shortparallel}{P_1}}{P_{n+1}}|+\dots+|\underset{\overset{\shortparallel}{x_n}}{x_{2n}}-\underset{\overset{\shortparallel}{P_n}}{P_{2n}}| =$$

$$=(|x_1-P_1|)^2+(|x_2-P_2|)^2+\dots+(|x_n-P_n|)^2=(x_1-P_1)^2+(x_2-P_2)^2+\dots+(x_n-P_n)^2$$

since all values bound to be ~~positive~~ non negative by the second constraint minimizing this problem is the same as minimizing $\|x-p\|_1$ in the previous section, because the square function is monotonicly increasing, just like the absolute value.

Part 2- General preview material

1\. a. $\frac{\partial f(x)}{\partial x_i}=1$

b. $\begin{pmatrix}1\\1\\ \vdots\\1\end{pmatrix}$ $n \times 1$ vector

c. $\frac{\partial f}{\partial u}=\langle \nabla f, u\rangle = \nabla f^T u = \frac{1}{\sqrt{n}}+\frac{1}{\sqrt{n}}+\dots+\frac{1}{\sqrt{n}}=\frac{n \cdot 1}{\sqrt{n}}=\sqrt{n}$

it does not depend on x, f has no squared components so any x_i does not appear in the gradient, hence the directional derivative does not depend on x.

d. $\nabla f(x)=\nabla a^T x = a^T$

2\. a. We want to find Q s.t $x^TQx = x_1^2 + 2x_2^2 + x_1x_2$

$$Q = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow x^TQx = (x_1a_{11} + x_2a_{21},\ x_1a_{12} + x_2a_{22}) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} =$$

$$= a_{11}x_1^2 + a_{21}x_2x_1 + a_{12}x_1x_2 + a_{22}x_2^2 \Rightarrow a_{11} = 1$$

$a_{22} = 2$

$$Q = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 2 \end{pmatrix}$$

$a_{21} + a_{12} = 1 \underset{\substack{\uparrow \\ Q \text{ is symetric}}}{\Rightarrow} a_{12} = a_{21} = 0.5$

b. $$\frac{\partial f}{\partial x_1} = 2x_1 + x_2$$

$$\frac{\partial f}{\partial x_2} = 4x_2 + x_1 \quad \Rightarrow \quad \nabla f = \begin{pmatrix} 2x_1 + x_2 \\ 4x_2 + x_1 \end{pmatrix}$$

c. assume we have x^TQx for some $Q \in \mathbb{R}_{n\times n}$ we can write x^TQx as following:

$$x^T \underbrace{\begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{nj}x_j \end{pmatrix}}_{Qx} = \sum_{i=1}^n \sum_{j=1}^n x_i a_{ij} x_j \quad **$$

** can be writen as:

$$\sum_{i=1}^n \left(a_{ii}x_i^2 + \sum_{j\neq i} x_i a_{ij} x_j\right)$$

Taking derivative with respect to some element k:

$$2a_{kk}x_k + \sum_{j\neq k} a_{kj}x_j + \sum_{j\neq k} x_j a_{jk} = \sum_{i=1}^n a_{ki}x_i + \sum_{i=1}^n x_i a_{ik}$$

$$\Rightarrow \nabla f(x) = \begin{pmatrix} \sum_{i=1}^n x_i a_{i1} + \sum_{i=1}^n a_{1i}x_i \\ \sum_{i=1}^n x_i a_{i2} + \sum_{i=1}^n a_{2i}x_i \\ \vdots \\ \sum_{i=1}^n x_i a_{in} + \sum_{i=1}^n a_{ni}x_i \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n x_i a_{i1} \\ \sum_{i=1}^n x_i a_{i2} \\ \vdots \\ \sum_{i=1}^n x_i a_{in} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^n a_{1i}x_i \\ \sum_{i=1}^n a_{2i}x_i \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i \end{pmatrix}$$

$(f(x) = x^TQx)$

$$= Q^Tx + Qx = (Q^T + Q)x \underset{\substack{\uparrow \\ Q \text{ is symetric}}}{=} 2Qx$$

d. $$\frac{\partial^2 f}{\partial x_1^2} = 2 = \frac{\partial}{\partial x_1} \circ \frac{\partial f}{\partial x_1}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4 = \frac{\partial}{\partial x_2} \circ \frac{\partial f}{\partial x_2} \quad \Rightarrow \quad \nabla^2 f = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1} \circ \frac{\partial f}{\partial x_2} = \frac{\partial}{\partial x_1}(4x_2 + x_1) = 1 = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

e. based on c, $\nabla f = 2Qx \Rightarrow \nabla^2 f = 2Q$

f. as I showed in c, the general form of ∇f is $(Q^T+Q)x$

The general form of $\nabla^2 f = Q^T + Q$

3. a. $l(t) = x + tu$

b. $f(l(t)) = f(x+tu) = g(t)$

c. $g'(t) = \frac{d}{dt}(g(t)) = \frac{d}{dt}(f(x+tu)) = \nabla f^T(x+tu)\cdot u = \nabla f^T \cdot u$

f. from 2 we have: $\nabla f = \begin{pmatrix} 2x_1 + x_2 \\ 4x_2 + x_1 \end{pmatrix}$

The hessian is:

$$\begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$$

d. $g''(t) = \frac{d^2}{dt^2}(f(x+tu)) = \frac{d}{dt}(g'(t)) = \frac{d}{dt}(\nabla f(x+tu)\cdot u) =$

$= u^T \nabla^2 f(x+tu)\cdot u = u^T H(f(x+tu))\cdot u$

e. The same as we had in first directional derivative; we want $t=0$:

$$\frac{\partial^2 f}{\partial u^2} = u^T H(f(x))\cdot u = g''(0)$$

f. continued:

$$\frac{\partial^2 f}{\partial u^2} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} \frac{3}{\sqrt{2}} & \frac{5}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{3}{2} + \frac{5}{2} = 4$$