Submission for HW 3

CS : 427 Mathematics for Data Science

Autumn 2020-21

K. Sai Anuroop, 170030035

Sunday, October 18, 2020.

Problem 1

Prove that a function $f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its hessian is positive semi-definite.

Answer Assuming that f is twice differentiable and domain of f is convex.

Sketch of the proof:

(1) We first prove it for the case of $f: \mathbb{R} \to \mathbb{R}$ and then

(2) We use the first-order condition for convexity, i.e., f is convex if and only if $\text{dom} f$ is convex and

$$f(y) \geq f(x) + \nabla f(x)^T (y-x) \text{ holds } \forall\, x, y \in \text{dom} f.$$

to prove the statement in problem.

---

(1) Let us assume $n=1$.

$\Rightarrow$) Suppose $f: \mathbb{R} \to \mathbb{R}$ is convex. Let $x, y \in \text{dom} f$ where $y > x$. By the first order condition

$f(y) \geq f(x) + f'(x)(y-x)$ and $f(y) \leq f(x) + f'(y)(y-x)$

This means $f(x) + f'(x)(y-x) \leq f(y) \leq f(x) + f'(y)(y-x)$

$$\Rightarrow f'(x)(y-x) \leq f(y) - f(x) \leq f'(y)(y-x)$$

So we get $f'(y)(y-x) \geq f'(x)(y-x)$

$$\Rightarrow \quad f'(y)(y-x) - f'(x)(y-x) \geq 0$$

$$\Rightarrow \quad \frac{f'(y)-f'(x)}{y-x} \geq 0 \qquad \text{(Note that } y>x\text{)}.$$

Taking $y \to x$,

$$\lim_{y\to x} \frac{f'(y)-f'(x)}{y-x} = f''(x)$$

So, $$\lim_{y\to x} \frac{f'(y)-f'(x)}{y-x} = f''(x) \geq 0.$$

Therefore $f''(x) \geq 0$.

($\Leftarrow$ suppose $f''(z) \geq 0 \ \forall z \in \operatorname{dom} f$.

Consider any two points $x, y \in \operatorname{dom} f$ where $x<y$.

Now consider $$\int_x^y f''(z)(y-z)\,dz$$

Note that $f''(z) \geq 0 \ \forall z \in \operatorname{dom} f$.

and $y-z \geq 0 \ \forall z \in [x,y]$.

So, $0 \leq \int_x^y f''(z)(y-z)\,dz$

We now apply $\boxed{\int u\,dv = uv - \int v\,du}$ rule to solve the above integral.

$$0 \le \int_x^y f''(z)(y-z)\,dz$$

$$= \int_x^y (y-z)\,d(f'(z))$$

$$= (y-z)f'(z)\Big|_x^y - \int_x^y f'(z)(-dz)$$

$$= -(y-x)f'(x) + \int_x^y f'(z)\,dz$$

$$= -(y-x)f'(x) + f(z)\Big|_x^y$$

$$= -(y-x)f'(x) + f(y) - f(x)$$

$$\Rightarrow \quad f(y) - f(x) - (y-x)f'(x) \geq 0$$

i.e.; $\quad f(y) \geq f(x) + f'(x)(y-x)$.

Therefore f is convex.

$\square$

(end of 1 proof).

2 Consider that a function is convex if and only if it is convex on all lines. In other words, the function $g(t) = f(x_0 + tv)$ is convex in t for all $x_0 \in \text{dom } f$ and all v.

Therefore, for any $n \geq 1$, f is convex if and only if $g''(t) = v^T \nabla^2 f(x_0 + tv) v \geq 0$

$\forall\, x_0 \in \text{dom} f$, $v \in \mathbb{R}^n$ and t such that $x_0 + tv \in \text{dom} f$

Here, it is necessary and sufficient that

$\nabla^2 f(x) \succeq 0 \quad \forall\, x \in \text{dom} f$.

Therefore, $f: \mathbb{R}^n \to \mathbb{R}$ is convex if and only if $\nabla^2 f(x) \succeq 0$ (Hessian of f is positive semi-definite). $\square$

(end of ② proof).

<u>Problem 2</u> When is the epigraph of a function a halfspace? When is the epigraph of a function a convex cone? When is the epigraph of a function a polyhedron?

<u>Answer</u> An epigraph of a function f, denoted by epi f ($f: \mathbb{R}^n \to \mathbb{R}$) is defined as follows:

$$\text{epi} f := \{(x, t) : x \in \mathbb{R}^n, t \in \mathbb{R}, f(x) \le t\}$$

strictly speaking

$$\boxed{\text{epi} f := \{(x, t) \mid x \in \text{dom} f, f(x) \le t\}}$$

A halfspace is a set of the form $\{x \mid a^T x \le b\}$ where $x \in \mathbb{R}^n$, $a (\neq 0) \in \mathbb{R}^n$, $b \in \mathbb{R}$.

Suppose epigraph is a cone.

$\Rightarrow \forall a > 0$, if $(x, t) \in \text{epi} f$ then $(ax, at) \in \text{epi} f$

Clearly, $(x, f(x)) \in \text{epi} f$.

So, $(ax, af(x)) \in \text{epi} f$.

$$\Rightarrow f(ax) \leq af(x) \quad -(i)$$

Similarly, $(ax, f(ax)) \in \text{epi} f$. So, if the epigraph is a cone, $(x, f(ax)/a) \in \text{epi} f$

i.e., $f(x) \leq f(ax)/a$. $-(ii)$

By combining (i) and (ii) we get

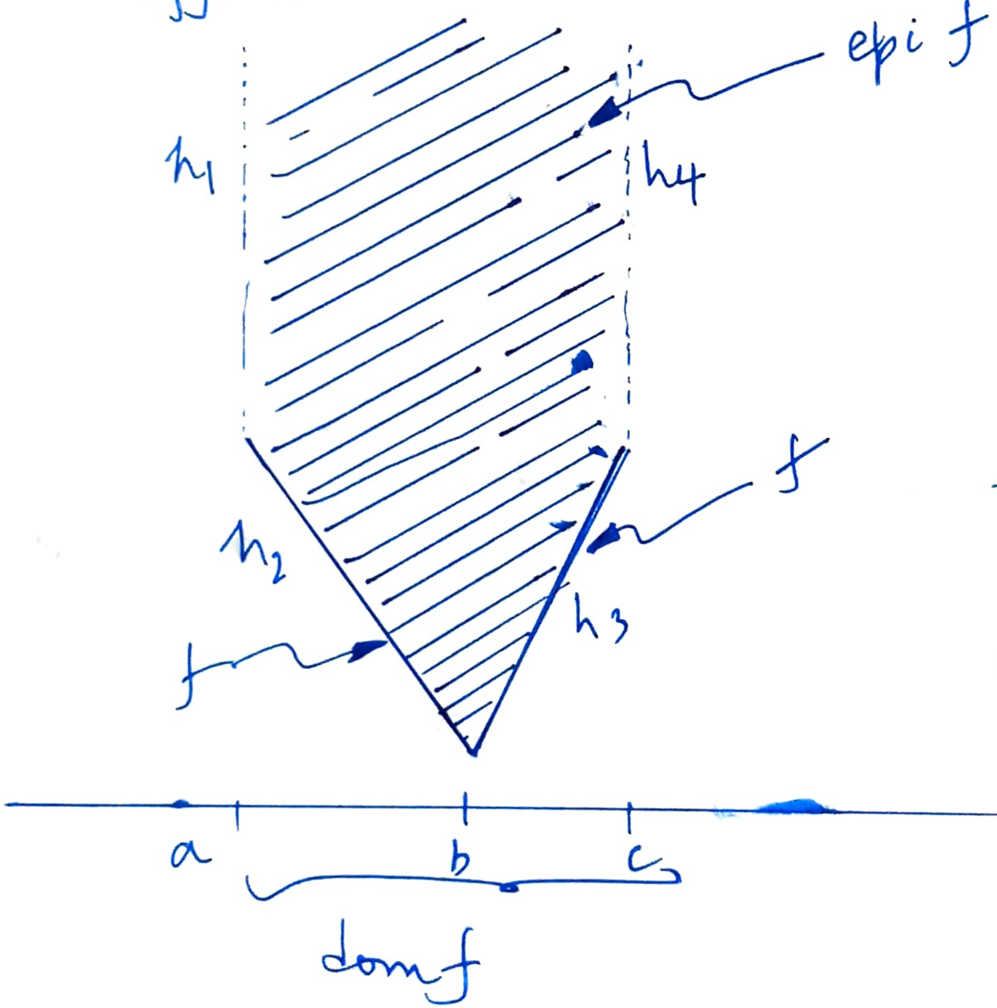
$$af(x) \leq f(ax) \leq af(x).$$

$\Rightarrow f(ax) = af(x)$. So this says f is positive homogenous.

$\therefore$ Epigraph of a function is a convex cone if f is ~~positively~~ homogenous.

A polyhedron is the intersection of a finite number of halfspaces and hyperplanes.

We just proved that the epigraph of f is a halfspace if f is affine. Thus for the epigraph to be a polyhedron, we require that f be 'piecewise affine'.



Note in this figure that f is piecewise affine. The shaded region is epi f which is a polyhedron too. The halfspaces which form epigraph are: h_1, h_2, h_3 & h_4 in the figure.

Problem 3 A function $f: \mathbb{R}^n \to \mathbb{R}^n$ is monotone if $\forall\, x, y \in \mathrm{dom}$

$$(f(x) - f(y))^T (x - y) \geq 0.$$

Show that gradient ∇f of f is monotone.
Is every monotone mapping a gradient of some convex function?

Answer The question is slightly incorrect. It should have been as follows:

[A function $f: \mathbb{R}^n \to \mathbb{R}^n$ is monotone if $\forall\, x, y \in \mathrm{dom}$

$$(f(x) - f(y))^T (x - y) \geq 0.$$

Suppose $g: \mathbb{R}^n \to \mathbb{R}$ is a differentiable convex function. Show that ∇g is monotone. Is every monotone mapping the gradient of some convex function?]

Answer

Since g is convex, we have

$$g(x) \geq g(y) + \nabla g(y)^T (x - y) \text{ and}$$
$$g(y) \geq g(x) + \nabla g(x)^T (y - x), \text{ for any } x, y \in \mathrm{dom} f$$

(by first order convexity).

Combining the above two inequalities we get

$$(\nabla g(x) - \nabla g(y))^T (x - y) \geq 0 \quad \text{which says}$$

∇g is monotone.

For the other part of the question, consider

$$\phi(x) = \begin{bmatrix} x_1 \\ x_1/2 + x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

<u>claim</u>

$\phi(x)$ is monotone.

$$\phi(x) = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\phi(y) = \begin{bmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\phi(x) - \phi(y) = \begin{bmatrix} 1 & 0 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix}$$

$$= \begin{bmatrix} x_1 - y_1 \\ \frac{x_1 - y_1}{2} + x_2 - y_2 \end{bmatrix}$$

$$(\phi(x) - \phi(y))^T (x - y) = \begin{bmatrix} x_1 - y_1 & \frac{x_1 - y_1}{2} + x_2 - y_2 \end{bmatrix} \begin{bmatrix} x_1 - y_1 \\ x_2 - y_2 \end{bmatrix}$$

$$= (x_1 - y_1)^2 + (x_2 - y_2)^2 + \frac{(x_2 - y_2)(x_1 - y_1)}{2}$$

$$\geq \frac{(x_1 - y_1)^2}{4} + \frac{(x_2 - y_2)^2}{4} + \frac{(x_2 - y_2)(x_1 - y_1)}{2}$$

$$= \left[\frac{(x_1 - y_1) + (x_2 - y_2)}{2}\right]^2$$

≥ 0. $\square$

(P.T.O)

† We know that for a given field $F = (F_1, \cdots, F_n)$ which is smooth (C^1) to be gradient, it is necessary that $\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$, $1 \leq i < j \leq n$.

Let us check if ϕ satisfies this condition.

$$\frac{\partial \phi_1}{\partial x_2} = \frac{\partial x_1}{\partial x_2} = 0$$

$$\frac{\partial \phi_2}{\partial x_1} = \frac{\partial \left(\frac{x_1}{2} + x_2\right)}{\partial x_1} = \frac{1}{2}$$

$$\Rightarrow \frac{\partial \phi_1}{\partial x_2} \neq \frac{\partial \phi_2}{\partial x_1}.$$

$\therefore$ ϕ cannot be a gradient.

So, we have shown a counter example where a monotone function is not the gradient of a convex function.

---

† This comes from the fact that curl of a gradient is 0

Problem 4 Suppose $f: \mathbb{R}^n \to \mathbb{R}$ is convex, $g: \mathbb{R}^n \to \mathbb{R}$ is concave, $\text{dom } f = \text{dom } g = \mathbb{R}^n$, and for all x, $g(x) \leq f(x)$. Show that there exists an affine function h such that for all x, $g(x) \leq h(x) \leq f(x)$.
In other words, if a concave function g is an underestimator of a convex function f, then we can fit an affine function between f and g.

Answer Recall that $\text{epi } f := \{(x,t) : x \in \text{dom } f, t \geq f(x)\}$

$$\text{hypo } f := \{(x,t) : x \in \text{dom } f, t \leq f(x)\}$$

Notice that $\text{int}(\text{epi } f) \neq \Phi$ (interior of epi f is not empty)

as $\text{dom } f = \mathbb{R}^n$.

Also, $\text{int}(\text{epi } f) \cap \text{hypo } g = \Phi$ as

$f(x) < t$ for $(x,t) \in \text{int epi } f$ and

$t \leq g(x)$ for $(x,t) \in \text{hypo } g$.

Thus, these two sets ($\text{int}(\text{epi } f)$ and $\text{hypo } g$) can be separated by a hyperplane. This means there exist $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ ~~($a \neq 0, b \neq 0$) su~~ (where both a and b are not zero at the same time), and $c \in \mathbb{R}$ such that

$$a^T x + bt \geq c \geq a^T y + bv \quad \text{where}$$

$t \geq f(x)$ and $v \leq g(y)$.

Note that if $b=0$, $a^T x \geq a^T y \Rightarrow a=0$.
$\forall x, y$

So, $b \neq 0$. Let $x=y$.

Then,
$$bt \geq c \geq bv.$$
$$\Rightarrow bt \geq bv. \quad -(i)$$

But we have $t > f(x)$ and $v \leq g(x)$

and $g(x) \leq f(x)$

So, $v \leq g(x) \leq f(x) < t$

$$\Rightarrow t > v. \quad -(ii)$$

Plugging (ii) in (i) we get,

$$b(t-v) \geq 0$$
$$\Rightarrow \underline{\underline{b > 0.}}$$

Let us now separate int(epi f) and hypo g by a hyperplane.

Consider a point ~~(x, t) ∈ int epi~~

$(x, t) \in \text{int}(\text{epi } f)$ and

$(y, v) = (x, g(x)) \in \text{hypo } g.$

Then $a^T x + bt \geq c \geq a^T x + b\,g(x)$.

Dividing by b. (note that we proved $b > 0$)

$$\frac{a^T x}{b} + t \geq \frac{c}{b} \geq \frac{a^T x}{b} + g(x)$$

$$\Rightarrow \boxed{t \geq \frac{-a^T x + c}{b} \geq g(x) \quad \forall\ t \geq f(x)}.$$

Thus, we have $h(x) = \dfrac{c - a^T x}{b}$, an affine function which lies between f and g.

A representation of the theorem is as follows: