Cache Ellsworth

Convex Analysis Exercises

Exercise 6.6

There are four critical points for this function.

1. $x = -\frac{1}{3},\ y = 0$
2. $x = 0,\ y = 0$
3. $x = 0,\ y = -\frac{1}{4}$
4. $x = -\frac{1}{9},\ y = -\frac{1}{12}$

Taking the second order derivative and solving for the eigenvalues, (Done in Python) We get the following eigenvalues

1. $\frac{1}{3},\ -3$
2. $1,\ -1$
3. $-2,\ \frac{1}{2}$
4. $-0.30854,\ -1.08034$

From these eigenvalues we see that the the following critical points are either local maxima, minima, or saddle points by the theory that if both eigenvalues are negative it is a local maxima and if one is is positive and one is negative it is a saddle point.

1. Saddle Point
2. Saddle Point
3. Saddle Point
4. Local Maxima

Exercise 6.7

i)

$$Q^T = (A^T + A)^T = A^{T^T} + A^T = A^T + A = Q.$$
$$\text{and } x^T Qx = x^T(A^T + A)x = x^T A^T x + x^T Ax.$$

We see that $(x^T Ax)^T = x^T A^{T^T} x = x^T Ax.x^T A^T X$ Is a scalar so it must be it's transpose is equal to itself. So then $x^T Qx = 2x^T Ax$.
ii) Using the GONC on $f(x)$. We assume x^* is a minimizer of F. This problem is unconstrained so x^* is

always an interior point. (Everything is feasible.). Then $Df(x^*) = 0$. So $Df(x^*) = \frac{1}{2}x^T(Q + Q^T) - b^T = 0$ (by Prop 6.4.6(ii)).

$$\implies \frac{1}{2}x^T(Q + Q^T) = b^T \implies \frac{1}{2}x^T 2Q = b^T \implies x^T Q = b^T \implies Q^T x = b.$$

iii)

Proof. ($\implies$:) We assume there's a minimizer solution, we call it x^*. Then by SONC and knowing it's an interior point we know that $D^2 f(x^*)$ is positive semidefinite. $D^2 f(x^*) = D(x^T Q - b^T) = Q^T$. So $Q^T \geq 0$. We assume by contradiction $Q^T = 0 \implies x^T Q^T x = 0$. Then $f(x) = -b^T x + c$ and, $f(x^*) = 0$. $f(x) > f(x^*)$ by assumption, so $-b^T x + c > c \implies -b^T x > 0 \implies b^T x < 0$. This is a contradiction because x can be infinitely big, and so can b because it is unconstrained. So $Q^T \neq 0$. Therefore $Q^T > 0$. So Q^T is semidefinite. ($\impliedby$:). We assume Q^T is semidefinite. By the invertible matrix theorem and because Q is symmetric, and a square, and all eigenvalues of Q not equal to zero. Then $Df(x) = Q^T x - b^T$ has a x_0 where $x_0^T Q = b$. This is a solution when $Df(x_0) = 0$. So therefore by $Df(x_0) = 0$ and $D^2 f(x_0) > 0$ we have a minimizer x_0 by the SOSC. Solving the system $Q^T x^* = b$ with positive definite Q is equivalent to solving the quadratic problem $f(x)$ because $Q^T x^* = b$ is just the FONC for the quadratic problem. □

Exercise 6.11

Proof. Let $f(x) = ax^2 + bx + c$ and $a > 0$ where $a, b, c \in \mathbb{R}$. Then Newton's Equation,

$$x_1 = x_0 - \frac{2x_0 a + b}{2a} = \frac{2ax_0 - 2x_0 a - b}{2a} = -\frac{b}{2a}$$

Therefore, this doesn't rely on any initial starting point. We know this is unique because f is quadratic. □

Exercise 6.15

See jupyter notebook Secant

Exercise 7.1

Proof. Let S be a nonempty subset of V. Conv(S) $= \{\lambda_1 x_1 + \cdots + \lambda_K x_K, x_i \in S, K \in \mathbb{N}\}$ We take $v, w \in Conv(S)$, so $v = \sum_k \lambda_{a,k} x_k$ where $\lambda_{a,k} \geq 0$ and $\sum_k \lambda_{a,k} = 1$. The w vector will have λ_b. We let $r = tv + (1-t)w$ where $\in [0,1]$. Then, we choose the maximum k of v and w. Then $r = \sum_k (t\lambda_{a,k} + (1-t)\lambda_{b,k}) x_k$. Then $\sum_k (t\lambda_{a,k} + (1-t)\lambda_{b,k}) = 1$ and $(t\lambda_{a,k} + (1-t)\lambda_{b,k}) \geq 0$. This implies that $r \in Conv(S)$. Thus, S is convex. □

Exercise 7.2

i)

Proof. Let P be a hyperplane in V. So $P = \{x \in V | \langle a, x \rangle = b, a \in V, a \neq 0, b \in \mathbb{R}\}$

$$\forall x, y \in P,\ \lambda x + (1-\lambda)y \in P,\ 0 \leq \lambda \leq 1$$

So we choose $x, y \in P$. Then we know that

$$a_1x_1 + \cdots + a_nx_n = b, a_1y_1 + \cdots + a_ny_n = b.. \text{ Then,}$$
$$\lambda(a_1x_1 + \cdots + a_nx_n) = \lambda b,\ (1-\lambda)(a_1y_1 + \cdots + a_ny_n) = (1-\lambda)b.$$
$$\text{adding the equations together, } \lambda(a_1x_1) + \cdots + \lambda(a_nx_n) + (1-\lambda)(a_1y_1) + \cdots + (1-\lambda)(a_ny_n) = b$$

So $\lambda x + (1-\lambda)y \in P$. ☐

ii)

Proof. Let H be a halfspace in V. So $H = \{v \in H | \langle a, v\rangle \leq b, a \in V, a \neq 0, b \in \mathbb{R}\}$

$$\forall v, w \in H,\ \lambda v + (1-\lambda)w \in H,\ 0 \leq \lambda \leq 1$$

So we choose $v, w \in H$. Then we know that

$$a_1v_1 + \cdots + a_nv_n \leq b, a_1w_1 + \cdots + a_nw_n \leq b.. \text{ Then,}$$
$$\lambda(a_1v_1 + \cdots + a_nv_n) \leq \lambda b,\ (1-\lambda)(a_1w_1 + \cdots + a_nw_n) \leq (1-\lambda)b.$$
$$\text{adding the equations together, } \lambda(a_1v_1) + \cdots + \lambda(a_nv_n) + (1-\lambda)(a_1w_1) + \cdots + (1-\lambda)(a_nw_n) \leq b$$

So $\lambda v + (1-\lambda)w \in H$. ☐

Exercise 7.4

i)

$$\begin{aligned}||x-y||^2 = ||x-y+p-p||^2 &= \langle x-p-y+p, x-p+p-y\rangle \\ &= \langle x-p, x-p\rangle + \langle p-y, p-y\rangle + 2\langle x-p, p-y\rangle \\ &= ||x-p||^2 + ||p-y||^2 + 2\langle x-p, p-y\rangle.\end{aligned}$$

ii) Let $\langle x-p, p-y\rangle \geq 0$. Thus by (i) and knowing that all norms are non-negative and that if $||p-y|| = 0$ then our assumed positive equation will not equal 0. Then,

$$||x-p||^2 < ||x-y||^2 \implies ||x-p|| < ||x-y||$$

iii) We apply (i).

$$\begin{aligned}||x-z||^2 &= ||x-p||^2 + ||p-\lambda y-(1-\lambda)p||^2 + 2\langle x-p, p-(\lambda y+(1-\lambda)p)\rangle \\ &= ||x-p||^2 + 2\lambda\langle x-p, p-y\rangle + \lambda^2||y-p||^2\end{aligned}$$

iv) Let p be a projection of x onto the convex set C and so $||x-p|| \leq ||x-y||$. When $\lambda = 1$, then $z = y$. Then $||x-z||^2 - ||x-p||^2 \geq 0.$. And by part (iii) and dividing by λ

$$0 \leq 2\langle x-p, p-y\rangle + \lambda||y-p||^2$$

Proof. ($\implies$:) Let P be a projection of x onto C. So $||x-p|| \leq ||x-y||$ and let $z = \lambda y + (1-\lambda)p$. Applying

(i), then (iii), then (iv), you get the result.
($\Longleftarrow$:) Use (ii) and then the definition of a projection. $\square$

Exercise 7.8

Proof. Let $f : \mathbb{R}^m \to \mathbb{R}$ is convex, $A \in M_{mxn}(\mathbb{R})$, and $b \in \mathbb{R}^m$. We define the function $g : \mathbb{R}^m \to \mathbb{R}$ as $g(x) = f(Ax + b)$. We let $x, y \in \mathbb{R}^n$ where $x \neq y$, and $\lambda \in [0, 1]$.

$$g(\lambda x + (1-\lambda)y) = f(A(\lambda x + (1-\lambda)y) + b) = f(\lambda A$$

$$f(\lambda(Ax + b) + (1-\lambda)(Ay + b)) \leq \lambda f(Ax + b) + (1-\lambda)f(Ay + b) = \lambda g(x) + (1-\lambda)g(y)$$

Therefore, g is convex. $\square$

Exercise 7.12

i)

Proof. Let $PD_n(\mathbb{R}$ be the set of positive-definite matrices in $M_n(\mathbb{R})$. We let $A, B \in PD_n(\mathbb{R})$ and $\lambda \in [0, 1]$. A, B are posititive-definite matrices. Then we know $0 < x^T Ax$ and $0 < x^T Bx$. Then,

$$x^T(\lambda A + (1-\lambda)B)x = \lambda(x^T Ax) + (1-\lambda)x^T Bx > 0$$

So this is a positive-definite matrix. Thus, this is a convex set. $\square$

ii) a) Let $g(t) = f(tA + (1-t)B)$ be convex. Therefore we know that with $(\lambda u + (1-\lambda)v)$ as our input for $g(\cdot)$:

$$\begin{aligned} f((\lambda u + (1-\lambda)v)A + (\lambda u + (1-\lambda)v)B) &= f(\lambda(uA + (1-u)B) + (1-\lambda)(vA + (1-v)B)) \\ &\leq \lambda f(uA + (1-u)B) + (1-\lambda)f(vA + (1-v)B). \end{aligned}$$

We let $X = uA + (1-u)B$ and $Y = vA + (1-v)B$. Therefore,

$$f(\lambda X + (1-\lambda)Y) \leq \lambda f(X) + (1-\lambda)f(Y)$$

Thus, f is convex.

b)

c) **
d)

$$g''(t) = \sum_{i=1}^{n} \frac{(1-\lambda_i)^2}{(t + (1-t)\lambda_i)^2}$$

These are both squared and so non-negative and so therefore the full equation must be non-negative.

Exercise 7.13

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ is convex and bounded above. So,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda b + (1-\lambda)b = b \forall x_1, x_2 \in \mathbb{R}^n$$
$$f(\cdot) < b \text{ for some } b \in \mathbb{R}$$

We assume f is not constant. Therefore $\exists x, y \in \mathbb{R}^n f(x) > f(y)$. Then,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \implies \frac{f(\lambda x_1 + (1-\lambda)x_2) - (1-\lambda)f(x_2)}{\lambda} \leq f(x_1)$$

Let $x = (\lambda x_1 + (1-\lambda)x_2) and y = x_2$, then

$$\frac{f(x) - (1-\lambda)f(y)}{\lambda} \leq f(\frac{x-(1-\lambda)y}{\lambda}) \implies \frac{f(x)-f(y)}{\lambda} + f(y) \leq f(\frac{x-(1-\lambda)y}{\lambda})$$

As $\lambda \to 0^+$ then $\frac{f(x)-f(y)}{\lambda} \to \infty$. Then f is not bounded. This is a contradiction. Then f must be constant. □

Exercise 7.20

Proof. Let $f : \mathbb{R}^n \to \mathbb{R}$ be convex and $-f$ also be convex. Then

$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \implies -f(\lambda x_1 + (1-\lambda)x_2) \geq -\lambda f(x_1) - (1-\lambda)f(x_2),$ also we know, $- f(\lambda x_1$

so then, $f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda)f(x_2)$. This is the definition of a linear transformation. So we know the linear transformation is a linear function. Thus, f is affine. □

Exercise 7.21

Let $D \subset \mathbb{R}$ with $f : \mathbb{R}^n \to D$ and $\phi : D \to \mathbb{R}$ is a strictly increasing function. Let $\mathcal{B}_\epsilon(x)$ be an open ball with radius ϵ.

($\implies$:). Assume x^* is a local minimizer for $\phi \circ f(x)$ s.t. $G(x) \preceq 0$, $H(x) = 0$. So $\phi(f(x^*)) \leq \phi(f(x)) \forall x \in \mathcal{B}_\epsilon(x)$. Then because ϕ is an increasing function then $f(x^*) \leq f(x) \forall x$. So x^* is a local min.

($\impliedby$:) Assume $x^* \in \mathbb{R}^n$ be a local min of $f(x)$. So $f(x^*) \leq f(x) \forall x \in \mathcal{B}_\epsilon(x)$. When we apply a strictly increasing function then $\phi f(x^*) \leq \phi f(x) \forall x \in \mathcal{B}_\epsilon$. So, x^* is a local min of $\phi \circ f$.