Problem Set #2

Ari Boyarsky
aboyarsky@uchicago.edu
OSM Bootcamp Math

July 11, 2018

Exercise 6.6

$$\nabla f(x,y) = (6xy + 4y^2 + y, 3x^2 + 8xy + x)$$

$$6xy + 4y^2 + y = 0 \implies y = 0 \text{ or } 6x + 4y + 1 = 0 \implies x = \frac{-4y-1}{6}$$

$$y = 0 \implies 3x^2 + x = 0 \implies x = 0 \text{ or } 3x + 1 = 0 \implies x = -1/3$$

$$x = \frac{-4y-1}{6} \implies 3(\frac{-4y-1}{6})^2 + 8(\frac{-4y-1}{6})y + \frac{-4y-1}{6} = 0 \implies y = -1/4 \text{ or } -1/12 \implies x = 0 \text{ or } -1/9$$

So, critical points are $(0,0), (0,-1/4), (-1/9,-1/12), (-1/3,0)$ The Hessian is,

$$D^2 f(x,y) = \begin{bmatrix} 6y & 6x+8y+1 \\ 6x+8y+1 & 8x \end{bmatrix}$$

Plugging in and evaluating the eigenvalues reveals that the max is $(-1/9,-1/12)$ while rest are saddle points.

Exercise 6.7

1. Q is clearly symmetric by definition since $A^T + A = A + A^T$.

   $$x^T Q x = x^T (A^T + A) x = 2x^T A x$$

   This clearly yields the result when plugging in and dividing by 2.

2. $f'(x) = 2x^T Ax - b^T = 0 \implies 2x^T Ax = b^T \implies Qx = b^T$ from the last part. This gives the result.

3. Notice that $Q^T x = b$ is satisfied only if Q is invertible. Then, this implies that it is either positive or negative definite since if a matrix is invertible the determinant is non-zero which implies there are no zero eigenvalues. Since, we are looking for a min, we know it is pos definite.

Exercise 6.11

$$x_1 = x_0 - f'(x_0)/f'(x_0)$$

$$x_1 = -b/2ax_0 \implies f'(x_1) = 2ax_0 + b = -b + b = 0$$

$$f''(x_1) = 2a > 0$$

Exercise 6.15

See attached .ipynb file.

Exercise 7.1

Take $y_1, y_2 \in Conv(S)$. Then, $\lambda y_1 + (1-\lambda) y_2 = \lambda_1(\sum_i \theta_i x_i) + \lambda_2(\sum_i \psi_i x_i) = \sum_i \lambda\theta_i x_i + (1-\lambda)\psi_i x_i \in Conv(S)$ since the convex hull is the set of all convex combinations of elements of S and any elements in Conv(S) is a convex combo of elements in S.

Exercise 7.2

1. $x, y \in H = \{x \in V | \langle a, x\rangle = b\} \implies \langle a, x\rangle = \langle a, y\rangle = b \implies \lambda\langle a, x\rangle + (1-\lambda)\langle a, y\rangle = b \implies \lambda x + (1-\lambda) y = b \ \forall \ \lambda \in (0, 1)$

2. $x, y \in \{x \in V | \langle a, x\rangle \leq b\} \implies \langle a, x\rangle,\ \langle a, y\rangle \leq b \implies \lambda\langle a, x\rangle + (1-\lambda)\langle a, y\rangle \leq b \implies \lambda x + (1-\lambda) y \leq b \in H \ \forall \ \lambda \in (0, 1)$

Exercise 7.4

1. $||x-y||^2 = ||x-p+p-y||^2 = \langle x-p+p-y, x-p+p-y\rangle = ||x-p||^2 + ||p-y||^2 + 2\langle x-p, x-y\rangle$

2. $||x-y|| = ||x-p|| \iff y = p \implies ||x-y|| > ||x-p|| \iff y \neq p$ and 7.14 holds.

3. $||x-z||^2 = ||x-p+p-z||^2 = \langle x-p+p-z, x-p+p-z\rangle = ||x-p||^2 + ||p-y||^2 + 2\langle x-p, p-z\rangle = ||x-p||^2 + ||p-\lambda y + (1-\lambda)p||^2 + 2\langle x-p, p-(\lambda y + (1-\lambda)p)\rangle$ which clearly yields the result.

4. 7.15 yields $0 \leq 2\langle x-p, p-y\rangle + \lambda||y-p||^2$ since from the previous part we have $||x-z|| > ||x-p||$

Exercise 7.8

$y_1 = Ax_1 + b, y_2 = Ax_2 + b$ then, $g(\lambda y_1 + (1-\lambda)y_2) = f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b)) \leq \lambda f(Ax_1 + b) + (1-\lambda) f(Ax_2 + b) = \lambda g(y_1) + (1-\lambda) g(y_2)$.

Exercise 7.12

1. Take $x, y \in PD_n(\mathbb{R})$ then, $\lambda(x) + (1-\lambda)y = z \in PD_n(\mathbb{R}) \ \forall \ \lambda \in (0, 1)$ since, $v^T z v = v^T(\lambda x + (1-\lambda)y)v \geq 0$ by def of pos definite which of course also implies that the convex combo is positive definite which implies the convexity of the set.

2. First, we know $g(t) = f(tA + (1-t)B)$ is convex from the previous exercise.

Exercise 7.13

Suppose by way of contradiction that f is not constant, then since f is bounded take $x, ys.t.f(x) > f(y)$ then $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) \implies \frac{\lambda f(x)+(1-\lambda)f(y)}{\lambda} = \frac{f(x)-f(y)}{\lambda} \leq f(\frac{\lambda x+(1-\lambda)y}{\lambda})$ and $f(x) > f(y) \implies f(x) - f(y)/\lambda \to \infty \lambda \to 0$ thus we do not have a bounded function so we must have a constant function.

Exercise 7.20

$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ and $-f(tx + (1-t)y) \leq -tf(x) - (1-t)f(y)$. But then, $-tf(x) - (1-t)f(y) \geq -f(tx + (1-t)y) \leq -tf(x) - (1-t)f(y)$ so instead we have equality then, $f(tx + (1-t)y) = tf(x) + (1-t)f(y)$. Then we have exactly the definition of affine.

Exercise 7.21

Let x^* be a local minimizer. Then, $f(x^*) \leq f(x)$. Since ϕ is monotonic, x^* also minimizes $\phi \circ f$. In the reverse, $\phi(f(x^*)) \leq \phi(f(x))$ for some neighborhood of x^*. Then since ϕ is monotonic we also have that x^* is a local minimizer of f.