Math Problem Set 4

Matthew Brown
OSM Boot Camp 2018

July 9, 2018

Problem 6.6.

$$Df(x,y) = \begin{bmatrix} 6xy + 4y^2 + y & 3x^2 + 8xy + x \end{bmatrix}$$
$$D^2 f(x,y) = \begin{bmatrix} 6y & 6x+1 \\ 6x+1 & 8x \end{bmatrix}$$

The critical points are $(0,0)$ and $(-\frac{1}{3}, -\frac{2}{3})$. $D^2 f(0,0)$ is indefinite, so the point is not a local extremum. $D^2 f(-\frac{1}{3}, -\frac{2}{3})$ is actually "negative definite" (I think?) so the point is a local maximum.

Problem 6.11. The minimizer of f is $-\frac{b}{2a}$. If we apply Newton's method to any $x_0 \in \mathbb{R}$,

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)} = x_0 - \frac{2ax_0 + b}{2a} = \frac{2ax_0 - (2a_x0 + b)}{2a} = -\frac{b}{2a}$$

Problem 7.1. *Proof.* Consider two elements $a, b \in C = \mathrm{Conv}(S)$ Then by the definition of C, we can represent a, b as linear combinations

$$a = \lambda_1 x_1 + ... + \lambda_n x_n$$
$$b = \Lambda_1 y_1 + ... + \Lambda_n y_m$$

where $\{x_i\}_{i=1}^n, \{y_i\}_{j=1}^m \in S$ and $\sum_{i=1}^n \lambda_i = \sum_{j=1}^m \Lambda_j = 1$ with all $\lambda_i, \lambda_j \in [0,1]$.
Now consider a combination $ka + (1-k)b$, with $k \in [0,1]$. We want to show that this is in C. We have

$$\begin{aligned} ka + (1-k)b &= k(\lambda_1 x_1 + ... + \lambda_n x_n) + (1-k)(\Lambda_1 y_1 + ... + \Lambda_n y_m) \\ &= k\lambda_1 x_1 + ... + k\lambda_n x_n + ... + (1-k)\Lambda_1 y_1 + ... + (1-k)\Lambda_m y_m \end{aligned}$$

Define

$$\{r_i\}_{i=1}^{m+n} = \{k\lambda_1, ...k\lambda_n, (1-k)\Lambda_1, ..., (1-k)\Lambda_m\}$$

Then

$$\sum_{i=1}^{m+n} r_i = k \sum_{i=1}^n x_i + (1-k) \sum_{j=1}^m y_j = k \cdot 1 + (1-k) \cdot 1 = 1$$

Also, $r_i \in [0,1]$ for each i. And we showed above that we can write $ka + (1-k)b$ as a linear combination of elements in S with coefficients $r_1, ...r_{m+n}$. so $ka + (1-k)b \in C$, and this shows that C is convex $\square$

Problem 7.2. *Proof.* Two Parts:

- (i) Suppose H is a hyperplane. Let $x, y \in H, \lambda \in [0,1]$. We know $\langle a, x\rangle = b$ and $\langle a, y\rangle = b$, so

  $$\langle a, \lambda x + (1-\lambda)y\rangle = \langle a, \lambda x\rangle + \langle a, (1-\lambda)y\rangle = \lambda b + (1-\lambda)b = b$$

  And so $\lambda x + (1-\lambda)y \in H$.

- (ii) Very similar to (i)...
  Suppose H is the half-plane $\{x \in \mathbb{R}^n | \langle a, x\rangle \leq b\}$. Let $x, y \in H, \lambda \in [0,1]$. We know $\langle a, x\rangle = c$ and $\langle a, y\rangle = d$ for some $c, d \leq b$, so

  $$\langle a, \lambda x + (1-\lambda)y\rangle = \langle a, \lambda x\rangle + \langle a, (1-\lambda)y\rangle = \lambda c + (1-\lambda)c \leq \lambda b + (1-\lambda)b = b$$

  And so $\lambda x + (1-\lambda)y \in H$.

□

Problem 7.4. *Proof.* I'll first show the four facts.

- (i)

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

- (ii) We can use the identity from (i). See that $||p-y||^2$ is always strictly positive for $y \neq p$, and if 7.14 holds then the term $\langle x-p, p-y\rangle$ is also nonnegative and the result follows.

- (iii) This will just be more manipulation of $\langle\rangle$ (RETURN TO THIS PROBLEM)

- (iv)

□

Problem 7.13. I'll argue by contradiction. Suppose that f is convex and bounded above, but f is not a constant function. Then there exist points $x_1, x_2 \in \mathbb{R}^n$ where $f(x_1) \neq f(x_2)$. Let M be the upper bound for f.
Suppose WLOG that $f(x_2) \geq f(x_1)$ Consider the line through these two points $L_{f(a),f(b)} = \{af(x_1) + bf(x_2) | a+b=1\}$. I can choose a^*, b^* so that $a^* f(x_1) + b^* f(x_2) > M$. See that $f(x_2)$ is on the line segment between $a^* f(x_1)$ and $b^* f(x_2)$, so it can be expressed as

$$f(x_2) = \lambda a^* f(x_1) + (1-\lambda)b^* f(x_2)$$

for some $\lambda \in [0,1]$.
Recall that $a^* x_1 + b^* x_2$ is in the domain of our function, so we can also think about ... ABORTED ATTEMPT - I HAVE INTUITION IN GRAPHS BUT CANNOT FORMALIZE

Problem 7.20. The first thing to note is that that

$$\begin{aligned} -f \text{ is convex} &\iff -f(\lambda x_1 + (1-\lambda)x_2) \le -(\lambda f(x_1) + (1-\lambda)f(x_2)) \\ &\iff f(\lambda x_1 + (1-\lambda)x_2) \ge (\lambda f(x_1) + (1-\lambda)f(x_2)) \end{aligned}$$

for $\lambda \in [0,1]$. And if we combine with the fact that f is convex we see that

$$f(\lambda x_1 + (1-\lambda)x_2) = \lambda f(x_1) + (1-\lambda)f(x_2)$$

for $\lambda \in [0,1]$, which is quite a handy fact. Indeed, this looks a LOT like the conditions we need for linearity - I just need a way to pass from λ to other scalars... Ugh. I'm just hitting a roadblock tonight, I guess... PROBLEM INCOMPLETE

Problem 7.21. *Proof.* I'll show both implications.

- $\Rightarrow$ Suppose x^* minimizes f. Then there exists an open neighborhood U of x^* such that for $x \in U, f(x*) < f(x)$. Because ϕ is increasing, this implies that for $x \in U$, $\phi(f(x^*)) < \phi(f(x))$ and x^* is a local minimizer for $\phi \circ f$.
- $\Leftarrow$ I'll show the contrapositive (even though that's probably needlessly complicating it, sorry). Suppose x^* does not minimize f. Then for any open neighborhood U of x^*, there exists $x_0 \in U$ such that $f(x*) \ge f(x_0)$. Because ϕ is increasing, this implies that for $\phi(f(x^*)) \ge \phi(f(x_0))$ and x^* is not a local minimizer for $\phi \circ f$.

$\square$