# ANALYSIS 2
# RECITATION SESSION OF WEEK 5

JACOB SHAPIRO

## 1. Exercise Sheet Number 5

1.1. **Convex Sets and Convex Functions.** Let $X$ be a vector space.

1.1. **Definition.** A convex set is a set $U \subseteq X$ such that if $(x_1, x_2) \in U^2$ then $[tx_1 + (1-t)x_2] \in U$ for all $t \in [0, 1]$.

The picture you should have in mind is that the straight line between each two points in $U$ is entirely inside of $U$.

1.2. **Example.** $\mathbb{R}^n$ is convex.

1.3. **Example.** A set in $\mathbb{R}^2$ which looks like a horse hoove is not convex.

1.4. **Definition.** Let $U$ be a convex subset of a vector space. Then a map $f : U \to \mathbb{R}$ is convex iff $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$ for all $t \in [0, 1]$ and $(x, y) \in U^2$.

1.5. **Example.** Let $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ be a norm on $\mathbb{R}^n$. Then $\|.\|$ is convex.

*Proof.* Let $t \in [0, 1]$ be given, and let $(x, y) \in [\mathbb{R}^n]^2$ be given. Then

$$\begin{aligned} \|tx + (1-t)y\| &\leqslant \|tx\| + \|(1-t)y\| \\ &= |t|\,\|x\| + |1-t|\,\|y\| \\ &= t\|x\| + (1-t)\|y\| \end{aligned}$$

□

1.6. **Example.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $x \mapsto x_1 x_2$. Then $f$ is not convex.

*Proof.* Let $t \in [0, 1]$ be given, and let $(x, y) \in \left[\mathbb{R}^2\right]^2$ be given. Then

$$\begin{aligned} f(tx + (1-t)y) &= f\left(\begin{bmatrix} tx_1 + (1-t)y_1 \\ tx_2 + (1-t)y_2 \end{bmatrix}\right) \\ &= (tx_1 + (1-t)y_1)(tx_2 + (1-t)y_2) \\ &= t^2 x_1 x_2 + (1-t)^2 y_1 y_2 + t(1-t)(x_1 y_2 + y_1 x_2) \\ &= t^2 f(x) + (1-t)^2 f(y) + t(1-t)(x_1 y_2 + y_1 x_2) \end{aligned}$$

Pick $x = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ and $y = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ then we have

$$f(tx + (1-t)y) = t(1-t)\,\forall t$$

whereas

$$tf(x) + (1-t)f(y) = 0 \forall t$$

and so any $t \in [0, 1]$ violates the convexivity condition. □

1.2. **Critical Points, Saddle Points Local Minima and Maxima.** Let $X$ and $Y$ be Banach spaces, $E \in \mathrm{Open}(X)$, and let $f \in C^3(E, Y)$.

1.7. **Definition.** A critical point of $f$ is a point $x_0 \in E$ such that either $\partial_v f(x_0) = 0\ \forall v \in X$ or $\nexists \partial_v f(x_0)$ for some $v \in X$.

(Think of $x \mapsto |x|$ when at $0$ this map is not differentiable).
Now assume $X$ and $Y$ are finite dimensional. Recall that we may approximate $f$ near an extremum point via

$$f(x_0 + x) \approx f(x_0) + \frac{1}{2}\langle x, H(x_0)x\rangle$$

1.8. *Claim.* If the Hessian matrix $(\partial_i \partial_j f)$ of $f$ is positive definite at an extremum point $x_0$ then $x_0$ is a local minimum. If the matrix is negative definite then $x_0$ is a local maximum. Otherwise it is a saddle point.

Observe that a matrix $M$ is positive definite iff $\langle v, Mv\rangle > 0$ for all vectors $v$. That means iff $v^T M v > 0$. If we write $M = PDP^{-1}$ where $D$ is diagonal, then, this is equivalent to requiring that $v^T D v > 0$ for all vectors $v$, which means

$$\sum_{i=1}^{n} (v_i)^2 (D)_{ii}$$

so that if all the entries of $D$ are positive, then we get a positive result no matter which $v$ we pick. If there are mixed signs then that is no longer the case.

---

*Date*: 15 March 2015.

1.9. **Exercise.** Let $f : \mathbb{R}^2 \to \mathbb{R}$ be given by $x \mapsto (x_1)^4 - 8(x_1)^2 + (x_2)^4 - 18(x_2)^2$. Find the extrema of $f$.

*Proof.* We first compute the Jacobian and matrix:

$$\partial_1 f(x) = 4x_1^3 - 16x_1 = 4x_1\left(x_1^2 - 4\right)$$

$$\partial_2 f(x) = 4x_2^3 - 36x_2 = 4x_2\left(x_2^2 - 9\right)$$

clearly these partial derivatives always exist, and so we need to find points where they are all zero: $\begin{cases} 4x_1\left(x_1^2-4\right) & = 0 \\ 4x_2\left(x_2^2-9\right) & = 0 \end{cases}$

and so the we have $\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \pm 3 \end{bmatrix}, \begin{bmatrix} \pm 2 \\ 0 \end{bmatrix}, \begin{bmatrix} \pm 2 \\ \pm 3 \end{bmatrix} \right\}$, all together nine points.

Now compute the second partial derivatives to be able to compute the Hessian matrix:

$$\partial_1^2 f(x) = 12x_1^2 - 16 = 4\left(3x_1^2 - 4\right)$$

$$\partial_2\partial_1 f(x) = 0$$

$$\partial_2^2 f(x) = 12x_2^2 - 36 = 12\left(x_2^2 - 3\right)$$

and so we have $H(x) = \begin{bmatrix} 12x_1^2 - 16 & 0 \\ 0 & 12x_2^2 - 36 \end{bmatrix}$ and at the extremum points we have:

(1) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$: $H\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} -16 & 0 \\ 0 & -36 \end{bmatrix}$ which is negative definite and so $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a local maximum.

(2) $\begin{bmatrix} 0 \\ \pm 3 \end{bmatrix}$: $H\left(\begin{bmatrix} 0 \\ \pm 3 \end{bmatrix}\right) = \begin{bmatrix} -16 & 0 \\ 0 & 72 \end{bmatrix}$ which is indefinite and so $\begin{bmatrix} 0 \\ \pm 3 \end{bmatrix}$ are saddle points.

(3) $\begin{bmatrix} \pm 2 \\ 0 \end{bmatrix}$: $H\left(\begin{bmatrix} \pm 2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 32 & 0 \\ 0 & -36 \end{bmatrix}$ which is indefinite and so $\begin{bmatrix} \pm 2 \\ 0 \end{bmatrix}$ are saddle points.

(4) $\begin{bmatrix} \pm 2 \\ \pm 3 \end{bmatrix}$: $H\left(\begin{bmatrix} \pm 2 \\ \pm 3 \end{bmatrix}\right) = \begin{bmatrix} 32 & 0 \\ 0 & 72 \end{bmatrix}$ which is positive definite and so $\begin{bmatrix} \pm 2 \\ \pm 3 \end{bmatrix}$ are local minima.

□

1.3. **Multi-Index Notation.** Let $n \in \mathbb{N}\backslash\{0\}$. Let $(\alpha, \beta) \in \left[[\mathbb{N}\cup\{0\}]^n\right]^2$ and $x \in \mathbb{R}^n$. Define

$$\|\alpha\| := \sum_{i=1}^{n} \alpha_i$$

$$\alpha! := \prod_{i=1}^{n} \alpha_i!$$

$$\binom{\alpha}{\beta} := \prod_{i=1}^{n} \binom{\alpha_i}{\beta_i}$$

$$\binom{\|\alpha\|}{\alpha} := \frac{\|\alpha\|!}{\prod_{i=1}^{n} \alpha_i!}$$

$$x^\alpha := \prod_{i=1}^{n} (x_i)^{\alpha_i}$$

$$\partial^\alpha := \prod_{i=1}^{n} \partial_i^{\alpha_i}$$

This makes certain notations much easier. For example, for Taylor approximations:

$$f(x+h) = \sum_{\alpha\in[\mathbb{N}\cup\{0\}]^n} \frac{1}{\alpha!}\left(\partial^\alpha f(x)\right) h^\alpha$$