Contents

1 Topics for First Midterm

1.1 Basic Definations

norm
see note basic:definations:norm

inner product
see note basic:definations:inner product

metric / distance functions
see note basic:definations:metric

open and closed subset of metric space
see note basic:definations:metricspace-openset and basic:definations:metricspace-closedset

1.2 Basic Theorems

EVT Extreme value theorem

IVT Intermediate value theorem

MVT Mean value theorem

IVFT Inverse Function Theorem

IPFT Implicit Function Theorem

1.3 Basic topology of metric spaces

closure
the set with all its limit points

interior
union of all open set contained in A

exterior
union of all open set disjoint from A

boundary
points that are neither interior nor exterior

limits $f : A \subseteq X \to Y$
$f(x) \to y_0$ as $x \to x_0$ if $\forall$ open $V \ni y_0$ $\exists$ open $U \ni x_0$ $[x \in U \cap A \wedge x \neq x_0 \to f(x) \in V]$

continuity f is cts at x_0 if x_0 is isolated point or $(\lim_{x\to x_0} f(x)) = f(x_0)$

Cauchy sequences A sequences $\langle x_i \rangle$ is Cauchy if
$\forall \varepsilon \exists N[n, m > N \implies d(x_m, x_n) < \varepsilon]$

completeness A metric space X is complete if every Cauchy sequences converge(to some point in X).

compact sets every open cover of X has a finite subcover

connected sets X cannot be divided into two disjoint nonempty closed/open/clopen sets.

relatively open sets p26 A is relatively open in $Y \subseteq X$ if $\exists$ open $U \subseteq X$ such that $A = U \cap Y$

Proposition $f : X \to Y$ is cts iff $\forall$ open $V \in Y$, $F^{-1}(V)$ is open in X. Similarly for closed.

Bolzano–Weierstrass property
A subset $E \in \mathbb{R}^n$ satisfies the BW property if every suquence has a convergent subsequence.

Bolzano–Weierstrass theorem $E \in \mathbb{R}^n$ satisfies the BW property iff E is closed and bounded.

Heine-Borel theorem $E \in \mathbb{R}^n$ is compact iff E is closed and bounded.

Application/(topological invariant) Suppose $f : X \to Y$ is continuous and X is compact then $f(X)$ is compact

Extreme value theorem Suppose $f : X \to \mathbb{R}$ is continuous and X is compact then $\exists x_0 \in X$ such that $f(x) \leq f(x_0) \forall x \in X$.

Path connected A set E is path connected if $\forall x, y \in E$, $\exists$ continuous map $f : [a, b] \to E$ such that $f(a) = x$ and $f(b) = y$.

Proposition If E is connected, and $f : E \to Y$ is continuous then $f(E)$ is connected

Proposition If E is path connected then E is connected.

Intermediate Value Theorem Suppose $E \in \mathbb{R}$ is connected and $f : E \to \mathbb{R}$ is continuous. Suppose $f(x) = a$ and $f(y) = b$ for some $x, y \in E$ and $a < b$. Then $\forall a < c < b \exists$ some $z \in E$ such that $f(z) = c$.

The ϵ-neighborhood theorem Let X be a compact subspace of $\mathbb{R}$; Let U be an open set of $\mathbb{R}^n$ contaioning X, Then there is an $\epsilon > 0$ such that the ϵ-neighborhood of

X is contained in U.

Cauchy-Schwarz inequality; all norms on a finite-dimensional vector space are equivalent; Bolzano Weierstrass theorem; Heine-Borel theorem; the con- tinuous image of a compact set is compact; the continuous image of a connected set is connected; intermediate value theorem; extreme value theorem. minima and maxima of continuous functions on compact sets

1.4 Differentiation

Derivative

- definition of the derivative
- partial derivatives
- directional derivatives

chain rule

- $(f \circ g) = (f' \circ g) \cdot g'$
- $\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx}$

continuity and differentiability

- differentiable implies continuity
- C^1 implies differentiable
- C^2 implies equality of mixed partial derivatives

Jacobian matrix

$$Jf = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \ldots & \ddots & \ldots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

continuously differentiable functions if the derivative exists and the derivative is continuous

higher order derivatives second derivative or higher

gradient $\vec{\nabla} f = \sum_i (D_i f) e_i$ (aka, $(\nabla f(x)) \cdot v = f'(x; v)$)

geometry of the Jacobian, the rows, the columns TODO

1.5 Max-min problems

Multi Index see note

Taylor's theorem The taylor series is useful at critical points because if a is a critical point of f and f is C^2 at a then

$$f(a+h) = f(a) + \frac{1}{2}\sum_{i,j=1}^{n} \partial_i\partial_j f(a)h^i h^j + R_{a,2}(h)$$

where

$$R_{a,2}(h) \to 0 \text{ as } x \to 0$$

Basic facts about the gradient $\nabla f(a)$ points in the direction of maximal increase and $|\nabla f(a)|$ is the rate of change of f in the direction of fastest increase. $\nabla f(a)$ is orthogonal to the level set of f that passes through a.

Critical points $a \in \mathbb{R}$ is said to be a critical point of f if $Df = 0$.

Proposition If f has a local maximum/minimum at a and f is differentiable at a then $Df(a) = 0$.

the Hessian $H(f) = (D_i D_j f(a))$ which has n eigenvalues counting multiplicity.

Hessian Matrix

$$H(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \cdots & \ddots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Minors A $k \times k$ principle minor of a matrix M is the matrix restricted to the first k rows and the first k columns.

Quadratic forms Completing the square.

Proposition A symmetric matrix h is positive definate if the determinant of all its principle $k \times k$ minors are positive.

Classification of critical points Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is c^2, a is a critical point. Then $H(f)$ has n eigenvalues (counting multiplicity).

- a is a local minimum of all eigenvalues are positive.
- a is a local maximum if all eigenvalues are negative.
- a is a saddle point if k eigenvalues are positive and $n - k$ are negative.
- a is a non-degenerate critical point if all eigenvalues are non-zero

Folland 2.82 Suppose f is of class C^2 on an open set in $\mathbb{R}^2$ containing the point a, and suppose $Df(a) = 0$. Let $\alpha = \partial_1^2 f(a)$, $\beta = \partial_1\partial_2 f(a)$, $\gamma = \partial_2^2 f(a)$. Then:

- If $\alpha\gamma - \beta^2 < 0$, f has a saddle point at a.
- If $\alpha\gamma - \beta^2 > 0$ and $\alpha > 0$, f has a local minimum at a.
- If $\alpha\gamma - \beta^2 > 0$ and $\alpha < 0$, f has a local maximum at a.
- If $\alpha\gamma - \beta^2 = 0$, no conclusion can be drawn.

Max-min problems with constraints Apply the classification of critical points for points in the interior. On the boundary use lagrange multipliers.

Lagrange multipliers Consider the functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g_1 : \mathbb{R}^n \to \mathbb{R}$ and $g_2 : \mathbb{R}^n \to \mathbb{R}$. We want to find the extreme values for f subject to $g_1 = c_1$ and $g_2 = c_2$. Then we just need find the values $a \in \mathbb{R}^n$ that satisfy the system of equations:

- $\nabla f(a) = \mu \nabla g_1(a) + \lambda \nabla g_2(a)$
- $g_1(a) = c_1$
- $g_2(a) = c_2$

1.6 Instructor's comment

General Comments: Should you memorize proofs of theorems? It is very hard to memorize all proofs of all theorems. In the long run, it is much more efficient, as well as useful and interesting, to first try to understand the proofs, and internalize the methods of proof, as well as possible; then to remember just an outline of the proof, or some key idea; roughly speaking, the minimum you would need to allow yourself to reconstruct the proof out of your base of general knowledge/understanding. Remember: it is important to know not simply whether something is true, but why it is true.

2 Basic knowledge

2.1 Definations

2.1.1 Linear algebra

norm A function $p : V \to \mathbb{R}$ such that
For all $a \in F$ and all $u, v \in V$

- $p(v) \geq 0 \wedge [p(v) = 0 \iff v = 0]$ (separates points)
- $p(av) = |a|p(v)$ (absolute homogeneity)
- $p(u + v) \leq p(u) + p(v)$ (triangle inequality)

inner product A function $\langle x, y \rangle : V \times V \to \mathbb{R}$ such that
For all $x, y, z \in V$ and $c \in \mathbb{F}$.

- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$
- $\langle cx, y \rangle = c\langle x, y \rangle = \langle x, cy \rangle$
- $\langle x, x \rangle > 0$ if $x \neq 0$

2.1.2 Topology

metric A function $d : X \times X \to \mathbb{R}$ such that
For all $x, y, z \in X$

- $d(x, y) \geq 0$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$

ϵ-neighborhood $\mathbf{U}(x; \epsilon) = \{y | d(x, y) < \epsilon\}$

open set in metric space A set $U \subseteq X$ is said to be open in X if
$\forall x \in U \exists \epsilon > 0[\mathbf{U}(x; \epsilon) \subseteq U]$
note that finite intersections and arbitrary unions of open set are open set

closed set in metric space A set contains all its limit point.
note that closed set is complement of open set in topology

2.1.3 etc

Measure theory - measure zero

Let $A \subseteq \mathbb{R}^n$. We say A has <u>measure zero</u> in $\mathbb{R}^n$ if for every $\epsilon > 0$, there is a covering $Q_1, Q_2, \ldots$ of A by countably many rectangles such that $\sum_{i=1}^{\infty} v(Q_i) < \epsilon$. If this inequality holds, we often say that the <u>total volume</u> of the rectangles $Q_1, Q_2, \ldots$ is less than ϵ.

oscillation

Given $a \in Q$ define $A_\delta = \{f(x) | x \in Q \land |x - a| < \delta\}$. Let $M_\delta(f) = \sup A_\delta$, and let $m_\delta(f) = \inf A_\delta$, define oscillation at f by $\mathrm{osc}(f; a) = \inf_{\delta>0}[M_\delta(f) - m_\delta(f)]$. f is cts at a iff $\mathrm{osc}(f; a) = 0$

2.2 Notations

2.2.1 Multi Index

- $\alpha = (\alpha_1, \ldots, \alpha_n)$
- $|\alpha| = \alpha_1 + \cdots + \alpha_n$
- $\alpha! = \alpha_1! \cdots \alpha_n!$
- $x^\alpha = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$
- $\partial^\alpha f = \partial_1^{\alpha_1} \partial_2^{\alpha_2} \cdots \partial_n^{\alpha_n} f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$

2.2.2 Derivative

$Df(a)$
: derivative of f at a

$f'(a; u)$
: directional derivative of f at a respect to vector u.

$D_j f(a)$
: j^{th}partial derivative of f at a.

f_i
: i^{th} component function of f.

$\vec{\nabla} g$
: gradient of g, $\vec{\nabla} g = \mathbf{grad} g = \sum_i (D_i g) e_i$

Jf

Jacobian matrix, $J_{ij} = D_j f_i(a)$

2.2.3 Abbreviations

cts Continuous

msr Measure

2.3 Theorems

Extreme value theorem

Suppose $f : X \to \mathbb{R}$ is continuous and X is compact,
then $\exists x_0 \in X$ such that $\forall x \in X. f(x) \leq f(x_0)$.

Intermediate Value Theorem

Suppose $E \in \mathbb{R}$ is connected and $f : E \to \mathbb{R}$ is continuous.
Suppose $f(x) = a$ and $f(y) = b$ for some $x, y \in E$ and $a < b$.
Then $\forall a < c < b \exists$ some $z \in E$ such that $f(z) = c$.

Mean Value Theorem

Suppose $\phi : [a, b] \to \mathbb{R}$ is

- continuous at each point of **closed** interval $[a, b]$
- differentiable at each point of **open** interval (a, b)

Then there exists a point $c \in (a, b)$ such that $\phi(b) - \phi(a) = \phi'(c)(b - a)$.

2.3.1 Theorem {Munkers-11.1}

1. If $B \subseteq A$ and A has measure zero in $\mathbb{R}^n$, the so does B.
2. Let A be the union of the collection of sets $A_1, A_2, \ldots$ If each A_i has measure zero, so does A.
3. A set A has measure zero in $\mathbb{R}^n$ if and only i

2.4 Taylor's theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^k on an open convex set S. If $a \in S$ and $a+h \in S$, then

$$f(a+h) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha + R_{a,k}(h),$$

If f is of class C^{k+1} on S, for some $c \in (0,1)$ we have

$$R_{a,k}(h) = \sum_{|\alpha| = k+1} \frac{\partial^\alpha f(a+ch)}{\alpha!} h^\alpha$$

3 Differentiation

3.1 Derivative

3.1.1 Differentiable

f is differentiable at a if there is an n by m matrix B such that

$$\frac{f(a+h) - f(a) - B \cdot h}{|h|} \to 0 \quad \text{as} \quad h \to 0$$

The matrix B is unique.

3.1.2 Directional derivative

Given $u \in \mathbb{R}^m$ which $u \neq 0$ define

$$f'(a;u) = \lim_{t \to 0} \frac{f(a+tu) - f(a)}{t}$$

Provide the limit exists.

3.1.3 Partial derivative

Define the j^{th} partial derivative of f at a to be the directional derivative of f at a with respect to the vector e_j, provide derivative exists.

$$D_j f(a) = \lim_{t \to 0} \frac{f(a+te_j) - f(a)}{t}$$

3.1.4 Continuously Differentiable

A function is C^1 if all of its partial derivatives are continous. A function is C^r if all of its partial derivatives are C^{r-1}.

3.2 Theorems

Munkers.5.1

If f is differentiable at a then all directional derivative of f at a exist and $f'(a;u) = Df(a) \cdot u$

Munkers.5.2

If f is differentiable at a then f is continuous at a.

Munkers.5.3

If f is differentiable at a then $Df(a) = [D_1 f(a) \quad D_2 f(a) \quad \cdots \quad D_m f(a)]$.

Munkers.5.4

a. [f is differentiable at a] $\Leftrightarrow \forall i$[$f_i$ is differentiable at a].

b. If f is differentiable at a, then its derivative is the n by m matrix whose i^{th} row is the derivative of the function f_i. $(Df(a))_i = Df_i(a)$

3.3 Continuously Differentiable Functions

Munkers 6.1

If $f : [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that $f(b) - f(a) = f'(c)(b - a)$.

Munkers 7.3

Let A be open in $\mathbb{R}^m$; let $f : A \to \mathbb{R}$ be differentiable on A. If A contains the line segment with end points a and $a+h$, then there is a point $c = a+th$ with $0 < t < 1$ of this line segment such that $f(a + h) - f(a) = (Df(c))h$.

Munkers 6.2

Let A be open in $\mathbb{R}^m$. Suppose that the partial derivative $D_i f_i(x)$ of the component function of f exists at each point x of A and are continuous on A. Then f is differentiable at each point of A.

Munkers 6.3

Let A be open in $\mathbb{R}^m$, let $f : A \to \mathbb{R}$ be a function of class $\mathbf{C}^2$. Then for each $a \in A$: $D_k D_j f(a) = D_j D_k f(a)$.

3.4 Chain Rule

Let $A \subset \mathbb{R}^m$. Let $B \subset \mathbb{R}^n$. Let $f : A \to \mathbb{R}^n$ and $g : B \to \mathbb{R}^p$, with $f(A) \subset B$. Suppose $f(a) = b$. If f is diferentiable at a and g is differentiable at b, then the composite function $g \circ f$ is differentiable at a. Furthermore,

$$D(g \circ f)(a) = Dg(b) \cdot Df(a)$$

3.5 Inverse Function Theorem

Let A be open in $\mathbb{R}^n$. Let $f : A \to \mathbb{R}^n$ be of class C^r.

IF $Df(x)$ is invertible at $a \in A$.

THEN There exists a neighborhood of a such that

- $f|_U$ is injective AND $f(U) = V$ open in $\mathbb{R}^n$
- the inverse function is of class C^r
- $f^{-1}(y) = [f'(f^{-1}(y))]^{-1}$

3.6 Implicit Function Theorem

Munkers 9.1

Let A be open in $\mathbb{R}^{k+n}$, B be open in $\mathbb{R}^k$.

Let $f : A \to \mathbb{R}^n$, $g : B \to \mathbb{R}^n$ be differentiable.

Write f in the form $f(x, y)$, for $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$.

IF $f(x, g(x)) = 0$ AND $\frac{\partial f}{\partial y}$ is invertible

THEN $Dg(x) = -\left[\frac{\partial f}{\partial y}(x, g(x))\right]^{-1} \cdot \frac{\partial f}{\partial x}(x, g(x))$

Suppose $f : A \to \mathbb{R}^n$ be of class C^r.

Write f in the form $f(x, y)$, for $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^n$.

IF $(a, b) \in A$ AND $f(a, b) = 0$ AND $\det \frac{\partial f}{\partial y}(a, b) \neq 0$

THEN There exists $B \in \mathbb{R}^k, a \in B$ and a unique $g : B \to \mathbb{R}^n$ such that $g(a) = b$ AND $\forall x \in B. f(x, g(x)) = 0$ AND g is C^r

4 Integration

4.1 Fundamental theorem of Calculus

- If f is continuous on $[a, b]$, and if $F(X) = \int_a^x f$ for $x \in [a, b]$, then $F'(x)$ exists and equals $f(x)$.
- If ...

5 Differential Calculus

5.1 Taylor's theorem

Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is of class C^k on an open convex set S. If $a \in S$ and $a + h \in S$, then

$$f(a+h) = \sum_{|\alpha| \leq k} \frac{\partial^\alpha f(a)}{\alpha!} h^\alpha + R_{a,k}(h),$$

If f is of class C^{k+1} on S, for some $c \in (0,1)$ we have

$$R_{a,k}(h) = \sum_{|\alpha| = k+1} \frac{\partial^\alpha f(a+ch)}{\alpha!} h^\alpha$$